UNIVERSITY OF MICHIGAN
MAR 18 1952
BUSINESS ADMINISTRATION LIBRARY

Sec. 8

Journal of the
AMERICAN STATISTICAL ASSOCIATION

MARCH 1952

Man as a Planning Animal ..... *Lowell J. Reed* 1

Some Applications of Statistics for Auditing ..... *John Neter* 6

Fertility Trends and Differentials in the United States ..... *Clyde V. Kiser* 25

Estimation for Sub-Sampling Designs Employing the County as a Primary Sampling Unit ..... *Emil H. Jebe* 49

Latent Structure Analysis and Its Relation to Factor Analysis. . *Bert F. Green, Jr.* 71

Measuring Magnitudes and Trends in the Production of Livestock and Meat ... . *A. V. Nordquist* 77

Reprints of Abstracts in Statistical Methodology 91

BOOK REVIEWS by Willard D. Arant, William G. Cochran, W. S. Connor, Jr., Boyd Harshbarger, M. G. Kendall, Henry B. McDaniel, Kenneth D. Roose, E. S. Shaw, M. Brewster Smith, Clair Wilcox and William W. Wright 92

PUBLICATIONS RECEIVED 111

VOLUME 47 NUMBER 257

American Statistical Association

Organized November 27, 1839, Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and others who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association.

Annual Dues

Regular membership	$ 8*	Member's subscription to Biometrics	4
Student membership	4	Contributing membership	25
Introductory membership (first dues payment of members under 30 years old)	4	Institutional membership, minimum	100

* Includes $5 for a year's subscription to the *Journal of the American Statistical Association*, and $1.00 for a year's subscription to *The American Statistician*.

Subscription rate, $8 per year. Prices for back issues available on request. A cumulative Index to Volumes 1–34, 1888–1939, and Annual Indexes thereafter may be obtained from the Secretary.

For further information about the Association and membership application forms, write the Secretary, 1108 16th Street, N. W., Washington 6, D.C.

The present Institutional Members are:

American Telephone & Telegraph Company
Arabian American Oil Company
The Atlantic Refining Company
Armour & Company
Armstrong Cork Company
Chrysler Corporation
Cities Service Company
The Columbia Gas System, Inc.
Deere & Company
The Dow Chemical Company
Dun & Bradstreet, Inc.
The Equitable Life Assurance Society of the U. S.
Ford Motor Company
General Electric Company
General Foods Corporation
General Motors Corporation
Industrial Surveys Company
International Business Machines
Kennecott Copper Corporation
The McBee Company
McCann-Erickson, Inc.
Metropolitan Life Insurance Company
National Analysts, Inc.
National Bank of Detroit
Northwest Bancorporation
Philadelphia Electric Company
Public Service Electric & Gas Company
Remington Rand, Inc.
Sears Roebuck & Company
Socony-Vacuum Oil Company, Inc.
Standard Oil Company (Indiana)
Standard Oil Company of New Jersey
Sun Oil Company
Western Electric Company, Inc.

Published Quarterly by the AMERICAN STATISTICAL ASSOCIATION

Publication Office: 450 Ahnaip Street, Menasha, Wisconsin. Editorial Office: 207 Haskell Hall, University of Chicago, Chicago 37, Illinois. Business Office: 1108 16th St. N.W., Washington 6, D.C. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, section 538, P. L. & R., authorized March 25, 1936. Entered as second class matter at the post office at Menasha, Wisconsin.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 47 MARCH 1952 NUMBER 257

ARTICLES

Man as a Planning Animal . . . . . LOWELL J. REED 1

Some Applications of Statistics for Auditing . . . . JOHN NETER 6

Fertility Trends and Differentials in the United States . . CLYDE V. KISER 25

Estimation for Sub-Sampling Designs Employing the County as a Primary Sampling Unit . . . . . EMIL H. JEBE 49

Latent Structure Analysis and Its Relation to Factor Analysis . . . . . BERT F. GREEN, JR. 71

Measuring Magnitudes and Trends in the Production of Livestock and Meat . . . . . A. V. NORDQUIST 77

Reprints of Abstracts in Statistical Methodology 91

BOOK REVIEWS

BUROS, OSCAR KRISEN, ed., *Statistical Methodology Reviews* 92

DIXON, WILFRID J. AND MASSEY, FRANK J., JR., *An Introduction to Statistical Analysis* . . . . . WILLIAM G. COCHRAN 93

QUENOUILLE, M. H., *Introductory Statistics* . . WILLIAM W. WRIGHT 94

NATIONAL BUREAU OF STANDARDS, *Tables to Facilitate Sequential t-Tests* . 96

PAYNE, STANLEY L., *The Art of Asking Questions* . . . . . M. BREWSTER SMITH 97

TINBERGEN, J., *Econometrics* . . . . . M. G. KENDALL 99

NUTTER, G. WARREN, *The Extent of Enterprise Monopoly in the United States, 1899–1939* . . . . . CLAIR WILCOX 100

CHANDLER, LESTER V., *Inflation in the United States, 1940–1948* . . . . . E. S. SHAW 101

CREAMER, DANIEL AND BERNSTEIN, MARTIN, *Behavior of Wage Rates during Business Cycles* . . . . . WILLARD D. ARANT 103

HAYES, DOUGLAS A., *Business Confidence and Business Activity* . . . . . KENNETH D. ROOSE 104

GINZBERG, ELI, GINSBERG, SOL W., AXELRAD, SIDNEY AND HERMA, JOHN L., *Occupational Choice, an Approach to a General Theory* . . . . . HENRY B. MCDANIEL 106

NELSON, CLARENCE W., *Use of Factorial Design in Industrial Relations Research* . . . . . W. S. CONNOR, JR. 108

SOUTHERN COOPERATIVE SERIES BULLETIN, *Studies of Sampling Techniques and Chemical Analysis of Vegetables* . . . BOYD HARSHBARGER 109

BUS. ADM.
LIBRARY
HA
1
.A5

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

The Editors welcome the submission of manuscripts for possible publication. They should be typewritten entirely double-spaced, including footnotes, and two copies should be sent to the Editor, W. Allen Wallis, 207 Haskell Hall, University of Chicago, Chicago 37. Books for review should be sent to the same address. Unsolicited book reviews are not accepted, but suggestions of titles for review are welcome.

EDITOR

W. ALLEN WALLIS, *University of Chicago*
ASSISTANT TO THE EDITOR: MARGARET A. LABADIE

ASSOCIATE EDITORS

ALBERT H. BOWKER
Stanford University
SOLOMON FABRICANT
New York University
HOWARD L. JONES
Illinois Bell Telephone Co.
WILLIAM G. MADOW
University of Illinois
FREDERICK MOSTELLER
Harvard University
C. ASHLEY WRIGHT
Standard Oil Company (N.J.)

ADVISORY PANEL OF FORMER EDITORS

WILLIAM G. COCHRAN (1945–50)
Johns Hopkins University
WILLIAM F. OGBURN (1920–1925)
University of Chicago
FRANK A. ROSS (1926–34, 41–45)
Thetford, Vermont
FREDERICK F. STEPHAN (1935–40)
Princeton University

Bus. Adm.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 257 MARCH 1952 *Volume 47*

MAN AS A PLANNING ANIMAL*

LOWELL J. REED
The Johns Hopkins University

THE concept of planning is very common among human beings. Man lives his life with his attention focused on the things of the moment, the events of the past, and his intentions as to the activities of the future. These intentions as to the future are commonly called plans, but in considering them it is necessary to distinguish those designs that are the almost involuntary response to preceding events from those that arise from deliberate thought and the weighing of relative values of different choices. It is not always easy to separate the deliberate from the involuntary decision. We may perhaps be aided in making a differentiation by pausing for a moment to look at the evidence that other animals do plan.

Our judgment as to whether other animals do make plans for the future is of necessity based on our observation of their acts and our interpretation of these acts. We speculate whether they were performed by instinct or came from a process that we would call thought if we saw such acts performed by man himself. We have the familiar instances of the harvesting and storing of food against winter needs on the part of a number of different animals. We tend to call it a response to instinct when we see squirrels gathering and storing nuts, whereas we are apt to consider that the human being collecting food for the winter is exhibiting conscious planning.

I recently had an opportunity to observe animals under natural conditions engaged in enterprises sufficiently complex to make me speculate as to the extent to which the thought process was involved. A family of beavers settled into the swimming pool at my farm. At first we were delighted with the opportunity to observe them, but as they began to cut not only the shade trees but the apple trees in the orchard, we took steps to dispossess them. I first raised the gate of the dam

* Presidential Address at the 111th Annual Meeting of the American Statistical Association, Boston, December 27, 1951.

thinking that with the pool gone they would leave. The next morning I found that about two-thirds of the way up the pool they had constructed a very respectable dam of their own so that the water in the part of the pool they were using was within 8 or 10 inches of its previous position. That day I tore their dam completely down and early in the evening I took up an observation post on the bank of the stream at one end of their dam site.

They furnished me several hours of entertainment. My first knowledge of their activity came from hearing a tree being cut in the vicinity. In a shorter time than I would have thought possible one of the pair came swimming silently down to my post with a small birch, 12 or 14 feet long, trailing behind him. His first move was to cut off a branch and carry it into the hole that formed their home. As he entered I could hear the pups whine at the approach of their evening meal. On his return to the outside he started on his own meal and although it may have little value as evidence of planning, you may be interested in his table manners. He first gathered the tip leaves and twigs of a small branch together with his front paws and fed this bundle into his mouth, tip first, his jaws going with a speed that made his teeth sound like a sewing machine. This rapid snipping of the leaves and twigs seemed to be all the chewing that he did. He then cut a branch about an inch in diameter so that he had a piece about twenty inches long, and taking it in his paws, he put it to his mouth as though he were going to play a flute. He then proceeded to eat the bark from the stick much as we would eat sweet corn from the cob. His appetite satisfied, he turned to the dam project and stripping off the branches, he took the birch pole that was about ten feet long and placed it diagonally across the stream with the butt resting against the bank. Then, seizing the upper end with his paws and teeth, he rammed the butt end about a foot into the bank. At this point he went back into the woods for another tree and while he was gone, the female came out for the first time. She seemed disturbed to find the dam gone, examined the stick that had been laid and apparently not satisfied with its position, she pulled it out and placed it about two feet upstream. She then went back into the hole. When the male returned, he noticed at once that the pole had been moved, and after a good deal of investigation and consideration he moved the stick back to its original position. When he had gone for his next load, the female came out and moved the pole back to the position she had selected. When the male returned, he seemed to bow to the inevitable and the dam building proceeded.

I have been asked if I was sure as to the sex, and if it might not have

been the male that had the last word. I have to admit that I have no direct evidence on this question, but the whole procedure seemed so human that I feel 95% confident that this judgment is statistically acceptable.

The subsequent steps in this seemingly designed experiment were to bring other trees, which were placed alternately on opposite banks and cross-tied; to bring smaller pieces which were snarled into the structure; and as the dam floated on top of the water, to roll good-sized cobble stones on top of the woven mass to weight it down; and finally, to mat mud and fine debris into it to make it hold back water. The entire evening's activity appeared to be a parental project designed to make suitable living conditions for the family.

When we look at man's activities in the field of planning, we find many of them very similar to those of other animals. Judged by action alone, they might as well be called response to instinct as thoughtful planning. On the other hand, we do know that man as an individual and in social aggregates does conscious planning for the future, and does attempt to carry these projects into action.

I would like to speak of one aspect of man's planning in an individual sense that seems to me to have bearing on his social planning. When I look forward to next summer and plan to go to my summer place with my wife, when I correspond with my children as to the possibility of their being there at certain times, there is no doubt or uncertainty in my mind as to the individuals for whom I am planning. There is no need for prediction or statistical estimation as to the persons concerned. There may be uncertainties as to the plans themselves, but the persons for whom the plans are being made are clearly in mind. When we come to social planning, however, this is not the case. Any planning for the future of the people of this country demands that we first make some prediction or estimation as to the number of people for whom we are planning. Thus population prediction becomes a primary element of social planning.

Another distinction between planning at an individual level and at a social level arises from the fact that the subject of statistics plays a very small part in planning for the individual but plays a major role in planning for society. It is the role of statistics in this planning about which I wish to make a few comments.

To lay down future programs of the type we are considering here, we need what we call facts. Most of these facts we attempt to obtain in statistical terms. For example, if we are planning a program of medical services, we begin by developing statistics with regard to such things

as medical resources, hospital facilities, by specified types, number of doctors and nurses available, number of people affected with illness by type and severity, the extent to which people have immunity to certain diseases, the degree to which by artificial means they can be immunized, and so on, ad infinitum. We are led from one issue to another as we develop in greater and greater detail our knowledge of physiology, disease, and medical activities. Since most of our information is expressed in statistical terms, specialized statistical techniques for handling the data have been developed and this has led the statistician into the design of experiments for acquiring the basic knowledge, and has indeed led to real specialization of statisticians themselves. Staying within the medical field and without attempting enumeration broadly throughout all areas of statistics, we now have specialists in hospital statistics, in quality control of laboratory processes, in bioassay, in the mathematics of biology, and so on. The degree of specialization within medical statistics is in fact somewhat similar to the degree of specialization of medicine itself. The extent of this is illustrated in Professor Cochran's experience when his daughter broke her arm. A call to a general practitioner met with the response that he did not handle children but suggested a pediatrician. The pediatrician when called said that he treated children, but did not set bones and suggested an orthopedist. The orthopedist said he did set bones and asked which arm it was, to which Professor Cochran, taking no chances asked, "Which arm do you set?" It seems at times that the statisticians are proceeding to as high a degree of specialization as this.

This detailed elaboration of statistics, a great deal of which is due to the advances of mathematical statistics, forms the basis upon which we are establishing a real statistical science. It is the type of activity which makes us proud of our profession, for although statistics is in its infancy in a scientific sense, it is certain to influence profoundly all future scientific thinking.

On the other hand, as is the case with all movements toward specialization, broad objectives have a tendency to be lost in the wealth of intensive detail. This is particularly noticeable when one examines many of the attempts at social planning. We find various groups taking the abundance of statistical detail within their specific field and developing complex social plans without regard to the relation of this specialty to the whole. Keeping still within my own field of medical and biological statistics, we find elaborate reports dealing with the statistics of medical services and facilities, doctor supply, nurse supply, and so forth, but we do not find attempts to integrate the proposed

programs with similar programs in industry, business, education, or public service, to determine whether or not the plans in the aggregate have any relationship to the number of human beings for whom we are planning. There is almost no contact between the ideas of these specialized plans and the estimates that have been made of the future population. It is in this sense that our social planning differs very sharply from personal planning, for as I mentioned before, in the case of personal planning we are always very sharply conscious of the individuals for whom the plans are being made.

There is little to be gained by planning from idealistic concepts, the number of nurses per 1,000 population, the number of school teachers per 1,000 school children, and so on, if these plans aggregate to a demand for more young women to carry them into effect than there are young women in the population.

We may ask where the responsibility rests for providing this integration of our statistical facts into a total picture. I should reply that it rests on the statistician. Just as we recognize the problem of analysis of man's observations as a process of statistical reasoning, so also should we view the problem of synthesis of knowledge into a complete picture.

I have directed my remarks on this point mainly to social planning with illustrations drawn from the medical service field. The ideas apply, however, to all fields of statistical application. It is impossible to achieve comprehensive economic or business planning within the framework of economic and business statistics alone and assume that the demographic issues have been taken care of by such simple indices as dollars or quantity of goods per person. The factors of population growth must be woven into the analysis as an integral part of the problem.

Analysis is always easier than synthesis. In any field of human activity it is simpler to obtain individual bits of knowledge than it is to put those bits together into a general pattern of human behavior. So also in the various fields of science, it is easier to obtain specific pieces of scientific observation than it is to put those pieces together to form a general scientific law. The process of reasoning involved in integration or synthesis is, however, statistical in character, and I should hope that the statisticians of the future would take synthesis as well as analysis as a part of their responsibility and would play a leading role not only in the planning of broad programs, but also in the development of scientific generalizations.

SOME APPLICATIONS OF STATISTICS FOR AUDITING

JOHN NETER*
Syracuse University

Some applications of statistical techniques in controlling clerical accuracy, in sampling accounting records, and in sampling physical property are studied, with stress placed upon the lessons learned from these applications. Their relevance to the auditing profession as well as to business management in general is considered.

IN AUDITING, extensive use is made of samples; but in basing decisions on these samples, little if any use is made of statistical techniques. In other areas of accounting, however, the application of sound statistical techniques to the interpretation of sample data is becoming increasingly frequent. The purpose of this paper is to describe a number of applications of statistics in the area of accounting, which are particularly relevant to the problems encountered in auditing.

Auditing consists of the examination of accounting records, vouchers, and other financial and legal records and documents of an organization to ascertain the accuracy and integrity of the accounting, in particular as it is reflected in the statements of financial condition and of income [1]. The examination may be performed for an organization by its own employees or by independent public accountants. In either event, heavy reliance is usually placed on sampling or test-checking techniques. Accounts receivable are verified by circularizing a selected number of them; inventories are generally test-checked; vouchers of cash disbursements may be examined, not for the entire accounting period, but only for a portion thereof. In view of the extensive use of sampling, it is surprising indeed that auditors have seldom employed statistical techniques to help reach conclusions about the state of the accounting records.

A consequence of the lack of application of statistical techniques is that standards of sound auditing procedures are primarily subjective. The auditor usually has no objective criterion, for example, as to how much test-checking is enough. It may well be that on the whole too much sampling is being carried on today by auditors. Some writing has been done on this subject, beginning with a paper by Carman [2], with others by Abrams [3], Berylson [4], Cranstoun [5], Herbert [6], Prytherch [7], Vance [8a, b, c], and Neter [9a, b]. This writing all suffers,

* The author is greatly indebted to the individuals and organizations who have contributed case material for this survey. Particularly valuable help was received from W. Edwards Deming, K. W. Halbert, Howard L. Jones, and E. T. Magruder.

however, from the fact that no actual applications have been studied in order to learn answers to such questions as these: What kinds of problems are likely to be encountered? What particular statistical techniques are most suitable for this area of application? What levels of risk would be economical as well as adequate?

Statistical techniques have been applied to at least three areas of relevance to auditing. These are:

(1) control of clerical accuracy;
(2) sampling accounting records;
(3) sampling physical property.

Methods of controlling clerical accuracy *as the work is being performed* are of considerable significance to the auditor because of his interest in the maintenance of a fairly high level of clerical accuracy in the accounting records. Statistical methods of controlling accuracy are particularly relevant to this interest.

Sampling accounting records to obtain an estimate of a certain characteristic is a common auditing procedure. For example, the auditor may sample payroll records in order to determine the extent of inaccuracies in the past year's vouchers. Other information, not now generally obtained by sampling, such as the age distribution of accounts receivable which are not recorded on punch cards, might be obtained by sampling the accounting records.

Sampling physical property occurs rather often in the verification of inventory by the auditor, less often in the verification of physical plant.

Control of Clerical Accuracy

One of the earliest uses of statistical techniques in controlling clerical accuracy was made in the Census Bureau. Deming and Geoffrey [10] report that sampling verification was used in the coding and punching of population and housing data for the 1940 census, in those stages where exact conformity with the enumeration was not required.

Every twentieth card in a housing census folio, which includes the schedules assigned to four enumerators, was verified after a randomized start if the puncher's accuracy qualified him for sampling verification. Otherwise, 100% checking was employed. The random start was designated to each verifier daily so that neither the puncher nor the verifier knew it in advance. In order that sampling verification might be confined to operators with reliable performance, tolerance limits were designated as follows, after considering the level of errors which could be allowed and the proportion of punchers who would qualify at any given level:

To qualify—"At least two of the last four weeks must show an average error rate of not more than 1 wrong card per 100 cards punched, and no week of the last four shall show an average of more than 2 wrong cards per 100 cards punched. (Weeks during which fewer than 2,000 cards were punched will not be counted.) In addition to the above, only one of the last four weeks may include a folio for which there were more than 3 wrong cards per 100 cards punched. (Folios of fewer than 300 cards will not be counted.)"

To disqualify—"A puncher will be dropped from sample verification if the average error rate for any week, determined from samples of her work, exceeds 3 wrong cards per 100 cards punched, or if it exceeds 2 wrong cards per 100 cards punched for each of two weeks out of the last four" ([10] p. 357).

For administrative convenience, an error in the punching operation was defined as a card with one or more incorrect punches. A logical alternative definition would have been an incorrect punch. The total number of punchings is usually not available in cases such as this, however, and relating the number of incorrect punchings to the cards punched would somewhat complicate the mathematical model.

Savings in direct labor cost due to substituting the sampling procedure for 100% checking were reported as $263,000 up to the time the paper was written. Indirect savings alone covered the cost of administering the sampling plan.

Individual control charts for each puncher who was within the administrative tolerance limits were used to discover the causes of excessive error rates [11]. These causes ranged from illegible folios to recent illness of the puncher. It is important to remember that sampling inspection was not applied to a puncher's work until he had given evidence of continuous high-quality performance. Thus the chief function of the sampling procedure was to determine whether the high-quality performance was continuing. If it was not, 100% checking would again be used.

Ballowe [12] reports that Alden's, Inc., a mail order business, began in the spring of 1945 to apply control chart techniques to filling customers' orders in one of the merchandise departments. After an item ordered by a customer has been picked from stock on the basis of its catalog number, size, color, and quantity, it is checked and placed on a conveyor belt to a gravity chute. At the chute, 100 work units of merchandise are selected at random several times a day and inspected. A certain number of error possibilities, including catalog number, size, color, quantity, or price, were set up for each work unit. A work unit is considered to have been incorrectly handled if an error is made in any of the specified error possibilities. The error rate is posted on control

charts as quickly as possible and remedial action instituted when out-of-control points appear.

Before control charts were used in this particular merchandise department, the error rate was 3 per 100 work units. Within two weeks after the introduction of statistical control techniques the error rate dropped to 1.65%, and at the time of writing in 1949 was about 0.7%. In the period from January 1946 to December 1949, the error rate for all merchandise departments was reduced by 58% ([12], pp. 12, 14).

In the fall of 1945, similar control chart techniques were introduced at Alden's in the general offices. Among the operations put under control were the following:

(1) Open envelope, remove contents, verify remittance, apply each impression to order blank ([12], p. 14).
(2) Read order to see whether any phase of transaction will not be handled in regular mail order process. If so, apply special rubber stamp, making abstracts on special requests, inquiries and complaints ([12], p. 18).

Here also, a tremendous reduction in the error rate was achieved during 1946–1949.

More recently, the credit department at Alden's introduced control chart techniques in posting-checking operations, credit approval, follow-up typing, and related activities. Filing, for example, is sampled by collecting duplicate stencil impressions, showing the customer's name and address, at the time the work is assigned to a clerk. After the papers are filed, a sample of 100 names is selected from the duplicate stencil impressions, and the files are examined to determine whether the items selected were filed correctly. In the credit department, as in the other departments, substantial reductions in error rates were achieved by the application of statistical control techniques.

For convenience, all sample sizes at Alden's are 100 work units. Up to six samples may be taken in a day. While the control charts are kept by departments and thus include the work of a number of employees, records are kept of the number and types of errors of each employee for use in corrective action. Since the most effective way to eliminate errors is to make them impossible, Alden's management attempts to do this as soon as error conditions are found to exist. For example, it may be found that the transcription of a figure, during which a transposition can occur, is not at all necessary. Only if the error possibility can not be removed is the emphasis turned to the reduction of the error ratio ([12], p. 18).

The Illinois Bell Telephone Company installed as early as July 1945 a group-sequential sampling plan for verifying clerical work in the ac-

counting department [13]. Howard L. Jones [14] has reported one of these applications which consisted of verifying the punching of Social Security numbers on tabulating machine cards. The purpose of this sampling plan was not primarily to detect and correct errors in the work already completed, but rather to minimize errors in the work currently being performed by promptly revealing conditions requiring remedial action. An error was defined as the incorrect punching of one or more digits of the Social Security number on the tabulating machine card ([14], p. 6). Samples were selected from the work of each individual when this was possible. For practical purposes, a consecutively-performed segment of work was treated as if it were a random sample ([14], pp. 7, 8).

A group-sequential sampling plan was used systematically and continuously. The acceptable error rate was set at 0.3%, the unacceptable error rate at 0.9%. The maximum risk of accepting unsatisfactory work was placed at 10%, of rejecting satisfactory work at 1%. The size of each sample group was determined by requiring that the acceptance and rejection numbers for cumulative samples increase by unity for each successive group. This requirement makes the administration of the plan more simple and also assures that the risks (or OC curve) for grouped sampling are the same or almost the same as for sampling by individual units [15]. Actually, the size of the group sample was slightly rounded from the theoretical size to a more convenient number. The continuous application of this group-sequential sampling plan was specified as follows:

> A sample of the first work is verified at once. If the work is acceptable, the next sample is taken and verified two hours later. If the work is still acceptable, the next verification is made after one day has elapsed; and similarly, the succeeding verifications are made one week and finally one month later. Rejection of the work leads to remedial action, and the verification interval reverts back to the beginning of the sequence.

Whether work already performed is to be verified if the sample verification leads to rejection depends upon the seriousness of not finding errors. The main objective here was appropriate remedial action ([14], p. 8). A distinction was made between systematic errors, likely to be due to faulty instruction, and accidental errors, likely to be the result of poor working conditions, poorly designed working papers, illness, fatigue, inexperience, and other similar factors. The nature of the remedial action depends, of course, on the cause of the errors.

The Bell System has carried on extensive tests to determine the applicability of sequential sampling plans for controlling the quality of clerical work, particularly that involved in rating toll tickets, i.e.,

pricing long distance calls. The results of these tests have been very satisfactory. In one of a series of tests in 1948, for example, the plan not only provided for control of clerical accuracy by remedial action when necessary, but it also located 56% of all errors made during this time by sampling only 12% of the work ([13], p. 1). On the basis of this experience, certain guides for setting up sampling plans for this type of work have been established [16]:

(1) In general, the work of an individual employee should constitute a universe.
(2) Setting the unacceptable quality level three times as high as the acceptable quality level and specifying the maximum risks of accepting unsatisfactory work and of rejecting satisfactory work at 0.10 provides for economical average sample sizes. The acceptable quality level should be set so that a substantial portion of employees can meet this requirement. In this connection, control charts might be used to advantage before making a decision as to what quality level can be reached by most employees suited for the work.
(3) If the size of the sample group is such that acceptance and rejection numbers increase by unity, administration of the plan is facilitated.
(4) A system of verification intervals, such as that suggested by Jones, should be established. The intervals may be specified in terms of time, number of assignments, or a combination of the two. The exact intervals to be used will depend on the nature of the work examined, the degree of control to be exercised, and the time and money available for sampling verification.

Single-sampling plans are also being used in the Bell System to verify the clerical work involved in rack sorting of tickets, which is the sorting of toll tickets by the two right-hand or left-hand digits of the telephone number. The purpose of this verification is not to replace 100% checking but rather to provide for control over the quality of the work of the individual clerks. Nevertheless, during a test period an examination of 28% of the tickets sorted disclosed 64% of all the errors made. On the basis of these trials, a number of useful guides were found:

(1) For this type of verification, single sampling appears to be the only practical type, as the taking of a sample is an intricate process.
(2) If the sample leads to the conclusion that the work is unsatisfactory, the remainder of the sorted tickets should be examined.
(3) A mis-sorted ticket constitutes an error.
(4) For this application it seems reasonable to design the plan so that

the errors in the work subsequent to sample verification will not exceed 0.5 per 1,000 tickets on the average.

(5) Single-sampling seems to be practical in this particular case only if the assignment consists of at least 1,000 tickets and if all pockets in the rack can be used.

(6) The sample should be selected by examining the contents of a number of pockets. The number of tickets subject to verification will determine the number of pockets to be selected. The suggestion that pockets which, on the basis of past experience, include most of the sorting errors should form part of the sample probably has the effect of reducing the maximum average percentage of error (or "average outgoing quality limit") below the stated requirement.

(7) The sampling verification procedure should be applied to the work of each individual clerk, and verification intervals, similar to those previously mentioned, should be established.

(8) Tables stating precisely which pockets are to be examined for various lot sizes may be prepared. One possible disadvantage of this suggestion is that the clerks might learn in advance which part of their work is to be sampled.

Many other instances of the application of statistical techniques to the control of clerical accuracy may be found in the Bell System companies which, together with a number of other firms and individuals, have pioneered in this development. Sequential sampling plans are being used in the verification of posting entries of workmen's time from work reports to labor distribution summaries. Magruder [17] of the Chesapeake and Potomac Telephone Companies reports, among a number of recent applications, the use of a continuous-sampling inspection technique in the verification of Western-Electric Company billings for items shipped direct from suppliers, and the use of sampling techniques in verifying daily work reports as to accuracy of accounting for material and labor.

The Standard Register Company applies a single-sampling plan to the control of accuracy of its sales-invoices. Previous 100% inspection was costly and did not detect all erroneous invoices. Shartle [18] reports that the use of sampling techniques reduced by 47% the time spent in verifying invoices and simultaneously maintained a satisfactory accuracy level. Samples are selected by a subjective random procedure from every group of invoices processed. Each group contains the work of several clerks. While sampling is not applied to the work of each clerk, a record of errors, by frequency and type, is kept for each clerk

for remedial purposes. An invoice is incorrect if it contains an erasure, strike-over, transposition of figures, omission, incorrect quantity, incorrect unit price, incorrect extension, incorrect total, etc. For a lot of invoices to be satisfactory, at least 99.25% of the individual invoices must be correct; a quality level of 98% or less is unsatisfactory. Risks of rejecting satisfactory work and of accepting unsatisfactory work are set at 5% [19]. The entire lot is verified if the number of incorrect invoices in the sample equals or exceeds the rejection number. In addition, a control chart is employed. Shartle [20] reports that out-of-control points have brought to light conditions such as improper or nonuniform training and improper placement of personnel.

An application of the control chart technique to controlling the accuracy of recording plane reservations is reported by Brinkman [21] of United Air Lines. About 10,000 incoming messages representing all space transactions for the U.S. are received daily in Denver where space for seats on United Air Lines is controlled. The phoned messages are penciled on incoming message slips. The phone wires are tapped three times a day and 200 consecutive messages are recorded each time. A carbon of the message slip is then checked against the recording. A message slip may be incorrect as to flight number, date, number of seats, stations involved, or even the recording of some item completely foreign to those which entered the conversation. Three-sigma control limits are used, the average being about 99.5 correct messages per 100 transcribed. It is not considered necessary to keep control charts separately for each operator, since out-of-control points are rare, but records of errors by type and incidence are kept for each employee. An important reason for the use of control charts in this case is to convince management that the personal element in telephone communication does not preclude accurate work. The control chart technique is also applied to the transcription of the message slip data to the space-control charts.

The applications cited illustrate the great diversity of circumstances in which statistical techniques have been applied successfully to the control of clerical accuracy. What are the conditions necessary to make these applications successful? Following Jones' [14] suggestions, five requirements may be listed:

(1) The purpose of the application of sampling techniques should not be to discover every error made but rather to establish control over the quality of clerical work. If the detection of all errors is necessary, sampling techniques are inapplicable.
(2) The work should be divisible into essentially similar units; in other words, the operation is repetitious.

(3) The volume or flow of the work should be large. The reason for this requirement is that clerical work does not lend itself readily to a quantitative measurement. Usually, it must be classified dichotomously, and this makes relatively large samples necessary for reasonable protection against incorrect decisions. For these relatively large samples to be economical, it is in turn necessary that the lot size be fairly large.

(4) Incorrect units should be clearly defined, so that each unit of work can be classified readily as correct or incorrect.

(5) The work should be completed at fairly frequent intervals. This permits frequent sampling, and thus remedial action can be taken promptly when it is necessary.

A number of further conclusions can be drawn from the examples presented here:

(6) The auditor's best assurance that the clerical accuracy of the accounting records is reasonably satisfactory after his examination is that statistical control over clerical accuracy was exercised in the first place. He should, therefore, encourage wherever possible the use of statistical techniques to control accuracy as clerical work is done.

(7) The plans so far developed generally have utilized either control chart techniques or the continuous, systematic application of acceptance sampling plans. The latter sometimes involves 100% verification if the number of errors in the sample equals or exceeds the rejection number, but not always. At any rate, little attempt seems to have been made so far to use continuous sampling plans, such as those of Dodge [30], [31], Dodge-Torrey [32] and Wald-Wolfowitz [33], which provide protection relating to the entire process. While it is true that these continuous plans involve a considerable amount of record keeping and require that labor for verification be available on a demand basis, it may be possible to develop other plans which overcome these objections and still provide protection relating to the entire process. This would appear to be a most desirable step because routine clerical work does generally constitute a continuous process. Also, it may be possible to reduce inspection costs while still maintaining required protection because the frequency of sampling in continuous plans will be governed by the quality of past performance. A major step in this direction seems to be the various verification intervals used, for instance, in the Bell System, which take past quality performance into account in determining the frequency and extent of sampling.

In that connection, Jones ([14], pp. 8, 9) has pointed out that with short verification intervals, which is assumed to be the equivalent of sampling the same infinite lot a number of times, the risk of accepting the work each time is less than the specified risk for a single sample. Hence he suggests that, when short verification intervals are used, the risk of accepting unsatisfactory work specified for the sampling plan may be increased somewhat. It may be added that the risk of rejecting satisfactory work at least once on the basis of several samples is greater than the risk of rejecting it on the basis of a single sample. Thus, this risk for a single sample should be made rather small for short verification intervals. The risk of rejecting satisfactory work was actually specified to be only 0.01 in the example which Jones discusses.

(8) On the basis of the variety of cases in which control over clerical accuracy was successfully achieved by applying statistical sampling techniques, it is reasonable to conclude that many further types of situations, both in business and in government, can be handled successfully by statistical control techniques [36].

Sampling Accounting Records

Cases available in this area are rather scarce, the application of statistical sampling techniques to accounting records being quite new. The auditor generally samples accounting records in order to determine the correctness of the recording of a transaction. Although the sampling plans described in this section have other objectives, they can, nevertheless, be adapted to the auditor's purpose.

Magruder [22] has reported an interesting application made by the Chesapeake and Potomac Telephone Company of Baltimore City. It is necessary to ascertain periodically the distribution of telephones by type of apparatus, of which there are six. The plant department maintains records showing the type of apparatus at each customer location. While a complete inventory of these records could be taken, samples provide the information more quickly and cheaply, and substantially as accurately ([22], p. 2).

First the universe of telephones was divided into three strata:

(1) Dial offices, where a subscriber line card shows the number of telephones by type of apparatus for each telephone number.
(2) Non-dial offices, where a subscriber line card shows the total number of telephones by type of apparatus for all the customers on 1-party, 2-party, 4-party, and rural lines.
(3) Private branch exchanges (PBX's), where records of the number of telephones by type of apparatus for each PBX extension line are generally located.

The subscriber line card or the PBX extension line was chosen as the sampling unit, since the universe was already enumerated in this manner. Sample sizes were then determined by imposing certain precision requirements which will not be discussed here. The selection of the sample proceeded as follows:

(1) In dial offices, every 144th subscriber line card was drawn, starting with a randomly chosen line number for each office.
(2) In manual offices, each 96th subscriber line card was drawn, starting with a randomly chosen line number for each office.
(3) The sample of PBX extension lines was chosen in three stages, using sampling with probability proportionate to size, simple random sampling, and systematic sampling, successively.

For each subscriber line card or PBX extension line selected, information as to number of telephones by type of apparatus was then obtained and by appropriate techniques combined into population estimates.

A method of evaluating the precision of the sample was incorporated into the sample design. It consists of the use of subsamples, originally suggested by Tukey ([23], p. 96). If a sufficient number of adequately large independent subsamples is used, each covering the entire universe, information may be obtained from them as to the precision of the subsample results, even though the selection of each subsample was not random.

To obtain the information as to the distribution of telephones by type of apparatus by complete examination would constitute a rather costly job. The following statement by Magruder is, therefore, especially significant:

> The reconciliation of plant quantities with the accounting records is one field where major savings are in prospect by the use of sampling. We have done enough sampling in this field to feel definitely assured of success. Continued research is necessary, however, to reach sample designs of improved efficiency. This involves the usual problems of definition of sample unit, possibility of stratification, selection of sample elements and precision computations [24].

The Chesapeake and Potomac Telephone Company has applied statistical sampling techniques to other accounting records in order, for example, to obtain a distribution of disconnected equipment by age bands, to audit the classification of troubles reported by subscribers, and to segregate the book cost of outside plant according to its usage for local, state, or interstate business. The sample estimate of the proportion of plant devoted to interstate business, for instance, had a margin of uncertainty of 1.6% at the 2-sigma level and was obtained

at only about one-tenth of the cost of a complete survey. Furthermore, Magruder declares, "the sheer size of a complete survey mitigates against intelligent and detailed scrutiny of records" which is possible when sampling is used. Therefore "we have reason to believe that the sample result is more precise than a complete survey" [24]. This aspect of sampling has also been observed in other types of sample surveys.

Jones [25] has reported an application of statistical sampling techniques to accounting records by the Illinois Bell Telephone Company. The particular information desired by the company was the mean and distribution of the number of local telephone calls, according to the various classes of service offered, as well as the mean telephone usage for all classes of service combined. This information is obtainable from the company's billing records. In setting up the sampling plan, stratification was employed both by central office areas and by classes of service. For purposes of determining optimum allocation of the sample, it was found that the standard deviations of the local message usage for each class of service are about the same for the different central office areas, but that there are important differences between classes of services. The dispersion is greater for business than for residence telephones and greater for individual than for party lines. Minimum requirements as to accuracy were set with respect to the means and distributions of the telephone usage for each class of service, as well as with respect to the mean for all classes combined. Selection of the sample was performed by mentally dividing a given file into as many more or less equal parts as the number of cards to be selected from it, and then picking a card from each part in haphazard fashion. This is somewhat of a systematic sample, but it has been found that the sample means in this application appear to be distributed about the same as the means of random samples ([25], p. 7). Incidentally, a sampling interval of 100 would be poor for listings of telephone numbers since an unusually large proportion of customers with heavy usage have telephone numbers ending in even hundreds. Unused telephone numbers are distributed irregularly in the actual card file, so these constitute less of a problem than if some other record were used in making the sample selection. All customers with private switchboards were included in the sample because their number is relatively small and the distribution of usage is severely skewed to the right; similarly all customers with more than one line whose usage exceeds 5,000 units during the month were included. Methods of improving the randomization of the samples for the remaining customers are being considered.

Sampling of accounting records has also been undertaken by governmental agencies. The Bureau of Old-Age and Survivors Insurance of the Social Security Administration for 12 years has been sampling its

universe of account numbers, which is approaching 100 million, in order to obtain up-to-date information on the characteristics of the insured population and the operations of its program ([34] p. 1). The Bureau's experience indicates that the most feasible type of sampling in this case is digital sampling; that is, selecting all accounts which have a certain digit or digits in given locations of the serial number. The device of maintaining a sample of sufficient size for tabulating detailed data and using smaller subsamples from this larger one for tabulating other data has proven itself flexible and economical ([34], pp. 13–14).

Another instance which may be cited in this trend to sampling accounting records is the Bureau of Public Assistance's suggestion to states with large caseloads of old age assistance, for example, to sample their accounting records in order to obtain an estimate of the distribution of assistance payments by amounts ([35], pp. 2–3).

These examples lead to the following general observations:

(1) In each case cited the body of accounting records sampled was large. If the universe of accounting records were small, the application of sampling techniques designed to achieve a reasonable assurance of accuracy would probably be uneconomical.

(2) In each example, the sample was of a recurring nature. Hence, if the changes in the universe of the accounting records were gradual over time, a sampling plan could be used repeatedly with periodic modifications.

(3) The physical state of the accounting records should be such that a sample can be selected with relative ease. Records in card files, whether use of consecutive numbering is made or not, were suitable in the cases cited [37].

(4) It is desirable to know some universe characteristics that could be estimated from the sample, and to compare the sample estimate with the known value. In the Chesapeake and Potomac sample, for instance, if the ratio of residence telephone stations to total telephone stations is known, one can compare it with the proportion of residence stations to total stations in the sample. Actually a great many such comparisons were made to provide additional assurances of the representativeness of the sample.

(5) The auditor often encounters bodies of accounting records which are large. His sampling is usually of a recurring nature. Hence it would appear that under those circumstances the establishment of statistical sampling plans would be desirable. The characteristics which he would study, generally those which pertain to the accuracy of the recording of a transaction, would not be the same as the characteristics verified in the examples discussed above. Neverthe-

less, he would encounter the same problems as to choice of sampling units, stratifications to be employed, selection of the sample and size of the sample as have been met and adequately solved in the cases cited. While undoubtedly many unique problems will confront the auditor in his particular applications of statistical sampling techniques to accounting records, the experience from the examples above should at the least encourage him to experiment with the application of statistical techniques to accounting records.

(6) These successful applications of statistical sampling techniques to accounting records seem to have significance far beyond the auditor. In both government and business, decisions have to be made quickly on the basis of information contained in voluminous records. Sampling often can provide such information quickly and to the necessary accuracy at reasonable cost. Certainly the lack of complete accuracy of sampling results should not prevent the application of these techniques. The cases cited previously are by no means the only ones in which statistical sampling techniques have been applied to accounting records. Applications made so far, nevertheless, represent only a small beginning on a large field of potential applications—a field which will probably be developed rather rapidly with the need for quick decisions and lack of manpower in business and in government today.

Sampling Physical Property

The auditor often samples inventories to verify the quantity and to ascertain the quality condition of the items. Plant and equipment is sampled more rarely by him. The experience so far obtained from the application of statistical sampling techniques to physical property should be of great interest to the auditor.

Several companies in the Bell System have used statistical sampling techniques in order to determine the current average physical condition of their telephone plant, which consists of a wide array of distinct classes ranging from central office equipment and trucks to aerial and underground cable. The information was needed by the regulatory commissions for rate-making purposes. Jones ([25], pp. 7–11) and Magruder [26] have reported on these applications; the latter will be cited here chiefly, although the problems encountered and the methods of solution in the two instances were similar.

The first problem to be faced in determining the physical condition of the Company's property was the method of specifying the state of physical deterioration. It was found that under field-inspection conditions, a maximum of five condition-grades was practicable. They were

defined as shown in the following table ([26], Chapter II, p. 1; Chapter III, p. 4):

Condition-Grade	*Physical Condition*	*Per Cent Value (Illustrative)*
A	New	95
B	Good	75
C	Fair	50
D	Poor	25
E	Worn out	10

The per cent values reflecting the extent of relative physical deterioration were set on the basis of knowledge and experience combined with judgment, and may vary from one class of property to another. As a practical matter, it was found to be best to define grades A, B, C, D, and E on the basis of easily distinguished physical characteristics, and then to assign per cent values to each condition-grade, basing these largely on the age-bands which correspond to each defined condition-grade [27].

In order that the sample results possess any degree of validity, uniformity of judgment on the part of inspectors is essential. It has been found that by a thorough training of the smallest practicable number of inspectors this uniformity of judgment can be achieved. Preliminary evidence indicates that on the average 8 out of 10 inspectors, after adequate training, will classify a given item in the same grade, the remaining two inspections splitting evenly between one grade higher and one grade lower. This split, because of its symmetry, does not affect the average physical condition reported ([26], Chapter 2, p. 2).

The importance of human errors in a survey of this nature cannot be stressed enough. In any survey there are a number of sources of error, among which are the sampling errors. In this particular application human errors are an especially important source of error unless the inspectors are first trained to develop uniformity in classifying the same physical property into the same condition-grade. Furthermore, the inspectors must not be required to examine so many units of property that they cannot adequately examine each unit to be inspected. Here, then, is a case which requires the economic balancing of sampling and human errors in order to get the most reliable survey results for a given budget expenditure (see, e.g., [23], ch. 2). Deming [28] states that an accurate determination of the current average physical condition of plant in this case can only be carried out by sampling methods, using very small samples. Larger samples would involve human errors far outweighing the sampling errors that were encountered in the inspections which are being reported here.

The precision of a sample required for submission to a public service

commission is probably greater than that necessary for other purposes. Even so, a precision range of less than $\pm 1.0\%$ for the average per cent condition of the property as a whole, with a 99.5% assurance, has been found practicable. Sample sizes were determined for each class of property on the basis of required precision. The sampling unit was chosen so that it can be readily enumerated from the property records and its location can be definitely determined from the information on the record. Furthermore, to the extent practicable, it is a unit that draws in other classes of property or that constitutes a relatively large part of the investment in its particular account. The unit "pole-location" draws in not only the pole itself, but such other property items as crossarms, anchors, aerial cable, and cable terminals.

It was found that the method of selecting sampling units which best met the requirements of a property valuation and which also was easy to apply was systematic subsampling. To sample pole-locations, for example, 10 independent subsamples were used.

Assignment of the inspection work was so arranged that each inspector would contribute about the same number of inspections to each of the ten subsamples, covering every section of the area. In that way, the sample design provided not only a measure of the precision of the subsamples by means of the ten independent subsamples, but it also provided evidence on the uniformity of judgment of the inspectors. Analysis of variance techniques were applied in testing these various aspects.

After the sampling units had been inspected, the average per cent condition for each class of property was computed. These averages were then combined into the average per cent condition of all property by using the amount of investment for each class as a weight.

Magruder [17] has reported another application of statistical sampling techniques to physical property which will only be mentioned here. Part of the telephone companies' coin box revenue for each month is uncollected at the end of the month, being in the coin boxes. To determine the amount of such revenue, a sample of coin boxes is taken by the Chesapeake and Potomac Telephone Companies each month.

Deming [29] presented a number of applications of sampling physical materials at the meeting of the International Statistical Institute in 1949. Lots of refined sugar imported into the United States have been sampled in order to estimate the quantity of sugar in the entire lot ([29], pp. 5, 6). Two-stage sampling has been applied to sample lots of domestic wool stored in warehouses in order to determine the percentage of clean wool in the lot. The primary unit was the bale, the secondary unit the core from the bale. By considering the cost of moving the

bale into position and the cost of boring the bale, sample sizes for primary and secondary sampling units were determined to yield estimates of required precision at the most economical cost possible under the conditions ([29], pp. 9–11).

In his recent book *Some Theory of Sampling* ([23], ch. 11), Deming devotes a chapter to the quarterly taking of inventories of tires held by dealers registered with the Office of Price Administration. Again, problems of stratification, optimum allocation of sample to achieve the required precision with a minimum sample size, and method of selecting the sample units arose. In addition to these, the problem of non-response entered the picture, a problem not present in any of the cases cited previously. That a serious problem may arise when non-response is possible may be seen from the fact that the average tires per dealer was 50 per cent higher in both December 1944 and March 1945 for the sample of non-respondent dealers of September 1944 than the average inventory for all other dealers. This illustration serves to point out that one cannot simply assume without some evidence that the non-respondents are similar to the respondents.

These examples permit a few conclusions:

(1) The techniques applied to sampling of physical property were the same statistical techniques applied to other types of sample surveys.

(2) Special problems may be encountered, such as the method of selecting the sample or defining the physical condition of property. Special problems, however, are always present in statistical work.

(3) The fact that statistical sampling techniques have been applied successfully to problems ranging from a nation-wide inventory of tires held by dealers to the evaluation of the physical condition of the property of a large telephone company indicates that the auditor could have a useful tool for his reconciliation of inventory and plant and equipment to the accounting records. To make practicable use of statistical sampling techniques, the magnitude of the inventory and plant and equipment in terms of sampling units will have to be fairly large. Thus the auditor should experiment at first with his larger clients in applying statistical sampling techniques to physical property as well as to accounting records. Afterwards, the auditor may extend these activities to smaller concerns to determine at what size or level the application of statistical sampling techniques becomes uneconomical.

(4) One of the cases illustrates the very basic problem of human errors in sample surveys. Such errors are important in auditing. The auditor, therefore, must study the relative magnitudes of human

and sampling errors in the audit of physical property as well as of accounting records so that an economic balance between the two can be reached. In that connection, he must remember that a more thorough and efficient scrutiny of property or records is possible when the number of items to be scrutinized is small than when it is large.

(5) Business and government could probably make more extensive use of the sampling of physical property in order to have quick and ready information on which decisions and control may be based.

REFERENCES

[1] Holmes, Arthur W., *Auditing* (Chicago: Richard D. Irwin, Inc., 1946), p. 1.

[2] Carman, Lewis A., "The efficacy of tests," *American Accountant*, 18 (1933), 360–6.

[3] Abrams, Jerome, "Sampling theory applied to test audits," *New York Certified Public Accountant*, 17 (1947), 645–52.

[4] Berylson, Kermit J., "Sampling and internal control on non-detailed monthly audits," *New York Certified Public Accountant*, 17 (1947), 660–3.

[5] Cranstoun, William D., "A new look at basic auditing techniques," *Journal of Accountancy*, 86 (1948), 274–83.

[6] Herbert, Leo, "Practical sampling for auditors," *New York Certified Public Accountant*, 17 (1947), 57–61.

[7] Prytherch, Robert H., "How much test checking is enough?," *Journal of Accountancy*, 74 (1942), 525–30.

[8a] Vance, Lawrence L., *Scientific Method for Auditing* (Berkeley: University of California Press, 1950).

[8b] ———, "Auditing uses of probabilities in selecting and interpreting test checks," *Journal of Accountancy*, 88 (1949), 214–7.

[8c] ———, "How much test checking is enough?," *Accounting Review*, 26 (1951), 22–30.

[9a] Neter, John, "An investigation of the usefulness of statistical sampling methods in auditing," *Journal of Accountancy*, 87 (1949), 390–8.

[9b] ———, "The application of statistical techniques in auditing procedures," *New York Certified Public Accountant*, 19 (1949), 345–50, 363.

[10] Deming, W. Edwards and Geoffrey, Leon, "On sample inspection in the processing of census returns," *Journal of the American Statistical Association*, 36 (1941), 351–60.

[11] Deming, W. Edwards, personal communication, 1951.

[12] Ballowe, James M., *Statistical Quality Control of Clerical and Manual Operations*, Reprint No. 10, Fifth Midwest Quality Control Conference, 1950.

[13] Jones, Howard L., "Group sequential plans for sampling clerical work," Summary of paper read at meeting of American Statistical Association, Cleveland, 1948.

[14] Jones, Howard L., "Sampling plans for verifying clerical work," *Industrial Quality Control*, 3, No. 4 (1947), 5–11.

[15] Wald, Abraham, *Sequential Analysis* (New York: John Wiley & Sons, Inc., 1947), p. 103.

[16] For this and the following case, certain background material has kindly

been furnished the writer by Mr. K. W. Halbert of American Telephone and Telegraph Company.

[17] The Chesapeake and Potomac Telephone Companies, *Summary of New Sampling Applications, October 1949 to August 1950.*

[18] Shartle, Richard B., "Quality control in the office," *Paperwork Simplification*, Number 16, 1949, 11–3.

[19] Shartle, Richard B., personal communication, 1951.

[20] Shartle, Richard B., personal communication, 1950.

[21] Brinkman, J. S., "United Air Lines speeds reservations," *Paperwork Simplification*, Number 16, 1949, 9–10; personal communication, 1950.

[22] Chesapeake and Potomac Telephone Company of Baltimore City, *Sample Design—Reconciliation of Continuing Property Records: Station Apparatus Account*, 1950.

[23] Deming, William Edwards, *Some Theory of Sampling* (New York: John Wiley & Sons, Inc., 1950).

[24] Magruder, E. T., personal communication, 1950.

[25] Jones, Howard L., "Design of samples for 'within company' analysis and control," *Business Application of Statistical Sampling Methods*, Proceedings of conference conducted by University of Illinois and Chicago Chapter, American Statistical Association, May 1950.

[26] Magruder, E. T., *Statistical Methods for Appraising Telephone Property*, The Chesapeake and Potomac Telephone Company, 1950.

[27] Jones, Howard L., personal communication, 1951.

[28] Deming, W. Edwards, personal communication, 1951.

[29] Deming, W. Edwards, On the *Sampling of Physical Materials*, Paper presented at the meeting of the International Statistical Institute held in Berne, 1949, mimeographed.

[30] Dodge, H. F., "A sampling inspection plan for continuous production," *Annals of Mathematical Statistics*, 14 (1943), 264–79.

[31] Dodge, H. F., "Sampling plans for continuous production," *Industrial Quality Control*, 4, No. 3 (1947), 5–9.

[32] Dodge, H. F. and Torrey, M. N., "Additional continuous sampling inspection plans," *Industrial Quality Control*, 7, No. 5 (1951), 7–12.

[33] Wald, A. and Wolfowitz, J., "Sampling inspection plans for continuous production which insure a prescribed limit on the outgoing quality," *Annals of Mathematical Statistics*, 16 (1945), 30–49.

[34] Mandel, Benjamin J., "Sampling in the old-age and survivors insurance statistical program," Paper read at 1950 Annual Meeting of the American Statistical Association, revised January 12, 1951.

[35] Perkins, Walter M., "Applications of sampling techniques in public assistance research," Paper read at 1950 Annual Meeting of the American Statistical Association.

[36] Murdock, Bennet B., "Some aspects of quality control as applied to clerical operations," *Industrial Quality Control*, 7, No. 6 (1951), 39–44, has recently described a program to control clerical operations at the Prudential Insurance Company.

[37] Rosander, A. C., Blythe, R. H., Jr., and Johnson, D. W., "Sampling 1949 corporation income tax returns," *Journal of the American Statistical Association*, 46 (1951), 233–41, describe a case where the accounting records were not so easily accessible.

FERTILITY TRENDS AND DIFFERENTIALS IN THE UNITED STATES*

CLYDE V. KISER
Milbank Memorial Fund

For the analysis of fertility trends the author draws upon official data and presents a preview of some of the materials in P. K. Whelpton's forthcoming monograph: *Cohort Fertility: Native White Women in the United States.* His discussion of trends in fertility differentials since 1940 is based mainly upon the Census Bureau's periodic releases giving fertility ratios among groups sampled in the Current Population Survey during the past few years.

MUCH interest has been aroused over the spectacular increase in the birth rate in this country during the past decade. This paper is concerned mainly with two broad questions: Does the increase in the birth rate connote an actual reversal in the long-time downward trend in average size of completed family or is it simply a temporary phenomenon associated with wartime and postwar conditions? In what manner have the increases in the birth rate varied among different elements of the population and what bearing does this have on the trend of fertility differentials? The data presented in this paper do not permit categorical answers to these questions. Whelpton's materials indicate beyond doubt that much of the increase in the crude birth rate is due to increases in marriages and first births and they make it seem improbable that crude birth rates of the 1947–1950 levels will be maintained very long. On the other hand, they also indicate the strong possibility of at least a slight and temporary reversal of the downward trend in average number of births per woman of completed fertility. This average may easily be somewhat higher for the cohorts of native-white women born during 1910–1914 and 1915–1919 than for those born during 1905–1909. As for differentials, samples drawn by the Census Bureau's Current Population Survey leave little doubt that increases since 1940 in fertility ratios (children under 5 per 1,000 married women of childbearing age) have been proportionately highest among groups previously characterized by lowest fertility. However, the consequent marked narrowing of group differences in fertility ratios since 1940 may have no more duration than a flash-flood increase in number of children under 5.

* The substance of this paper was presented before a joint meeting of the American Statistical Association and the Population Association of America in Chicago on December 27, 1950.

FERTILITY TRENDS

A brief historical setting is essential for interpretation of recent fertility trends. Like other highly modernized countries, the United States has had a long history of declines in fertility. In fact, this country may provide an exception to the usual tendency for declines in mortality to precede declines in fertility. Whatever that situation may be, there is good evidence that fertility ratios began declining in this country at least by the second decade of the last century. Census data for the 1810–1940 period indicate a decline in the number of children under 5 per 1,000 white females 20–44 years of age with each successive census except one (that of 1860) and that exception may not be real. At all events, the fertility ratio of about 1,358 in 1810 was reduced by approximately one-third by 1850, by about one-half by 1900, and by about two-thirds by 1940, when it was 419.[1] However, a sharp reversal is evident from the preliminary age distributions for the 1950 Census, which yield a fertility ratio of about 561 for whites with no adjustment for underenumeration of children.[2]

In terms of birth rates, Thompson and Whelpton have estimated that in 1800 and 1810 there were approximately 55 births per 1,000 white population, about 43 in 1850, and about 30 in 1900.[3] During the period of the expanding birth registration area, birth rates for whites and nonwhites combined, corrected for underregistration and adjusted to the country as a whole, declined from 29.5 in 1915 to a low point of 18.4 in 1933. During this period the chief interruption of the decline came in 1920 and 1921 following the return of World War I soldiers in 1919. During the remainder of the 'thirties there was a slow and irregular increase to 19.2 in 1938 and 18.8 in 1939.

After 1940, and particularly after 1945, the upward trend of the birth rate was little short of phenomenal. The crude rate (adjusted for underregistration) rose from 19.4 in 1940 to a wartime high of 22.9 in 1943. Then with millions of men overseas, the birth rate declined to 20.7 by 1945. Following demobilization, the birth rate jumped again and in 1947 it reached 27.0, its highest point since 1921. The rate then dropped to 25.4 in 1948 and 25.1 in 1949.[4]

---

[1] For the complete series of fertility ratios for whites from 1810 to 1940, increased 5 per cent for underenumeration of children, and (for censuses after 1820) standardized for age of women, *see:*

Whelpton, P. K. *et al.*, *Forecasts of the Population of the United States, 1945–1975.* (Bureau of the Census, Washington: Government Printing Office, 1947), p. 16.

[2] U. S. Bureau of the Census: *1950 Census of Population, Preliminary Reports.* General Characteristics of the Population of the United States, April 1, 1950, Series PC-7, No. 1, February 25, 1951, pp. 6–8.

[3] Thompson, Warren S. and Whelpton, P. K., *Population Trends in the United States* (New York: McGraw-Hill Book Company, 1933), p. 263.

[4] National Office of Vital Statistics:

Although the relation of birth rates to marriage rates is by no means simple, the major dips and peaks of the annual birth rate have, with a lag of about one year, tended to follow those of the marriage rate during the past 20 years. This may be due mainly to a close relation of marriages in one year to first births the next year.[5] Owing to the economic depression, the marriages occurring in 1932 dropped to 7.9 per 1,000 population, the lowest rate on record for this country. This was followed by an all-time low in the birth rate in 1933. The increases in employment associated with defense preparations in 1940–1941 were accompanied by sharp increases in the marriage and birth rates. Pearl Harbor and our entrance into the war provided an added stimulus and December 1941 brought a sharp rise in the marriage rate. As already noted, the wartime birth rates reached their peak in 1943 in so far as complete years are concerned. On a monthly basis a wartime peak of 24.9 was reached in September and October 1942, nine and ten months after Pearl Harbor. After the war the marriage rate reached 16.4 per 1,000 population in 1946, the highest on record for this country, and this was followed by the postwar peak in the birth rate in 1947. On a monthly basis the postwar peak in the marriage rate came in June, 1946 and that of the birth rate in December, 1946.[6]

Demographers and many others are keenly interested in the significance of the wartime and postwar increases in the birth rate. P. K. Whelpton has studied this problem more intensively than has anyone else in the United States. His first analysis, published in 1945, served at once to point up the important bearing of marriages and first births on increases in fertility during the early 1940's and an important inadequacy of the net reproduction rate. To quote from this article: "With 1942 came the impossible. In a cohort having 1942 age-specific rates for first births, 109 of each 100 women living through

---

Vital Statistics of the United States, 1948, Part I (Washington: Government Printing Office, 1950), p. xix (Table Y).

Births and Birth Rates in the Entire United States, 1909–1948. *Vital Statistics—Special Reports*, Vol. 33, No. 8 (September 29, 1950), p. 141.

Summary of Natality Statistics, United States, 1949. *Vital Statistics—Special Reports*, Vol. 36, No. 1 (May 14, 1951), p. 4.

[5] A reader of this paper has pointed out (a) that the same factors stimulating marriages may also cause couples already married to have a child, and (b) that on a monthly basis, adjusted for seasonality, there are some wartime and postwar peaks due to second and third child births two to three years after a peak in the marriage rate.

[6] The following recent report from the National Office of Vital Statistics is also of interest: "In April, 1951 the estimated number of live births increased sharply over the figure for April, 1950 (13.1 per cent). This was the greatest rise in the monthly number of births from one year to the next since 1947. It follows by 9–10 months the beginning of the upswing in marriage licenses in the middle of 1950 which was coincident with the outbreak of hostilities in Korea." *Monthly Vital Statistics Bulletin*, Vol. 14, No. 4 (June 20, 1951), p. 1.

the childbearing period would have at least one child."[7] A later article showing trends in the number of births to native-white women during 1920–1947, by order of birth, brings out clearly the fact that the depression dip and the wartime and postwar increases in the births were sharpest for lower orders of birth.[8] Computations based upon births per 1,000 females 15–44, by order of birth, reported by the National Office of Vital Statistics, indicate that the percentage increase from 1940 to 1947 was highest for first births and became successively lower with increasing order of births through the seventh. The increases in this fertility rate were 67 per cent for first births, 59 per cent for second births, 48 per cent for third births, 29 per cent for fourth births, 16 per cent for fifth births, 2 per cent for sixth and seventh combined, and a decrease for births of eighth and higher orders. It should also be noted, however, that the 1947–1949 *decrease* in the birth rate was due mainly to a decrease in first births. There were slight increases in the rates of second and higher order births from 1947 to 1949.[9]

The most thorough-going analysis of fertility trends in this country is afforded in Whelpton's recently completed study, *Cohort Fertility: Native White Women in the United States*, being published by the Princeton University Press. The writer is indebted to Professor Whelpton for the privilege of examining his manuscript and for permission to give a preview of certain data included and several of the outstanding conclusions.

In scope and detail of data and in potential value to demographers of the future, this latest work of Whelpton stands easily on a par with T. H. C. Stevenson's monumental report on the fertility of British marriages.[10] The manuscript consists of about 400 typed pages, about half of which are fertility tables for cohorts. There are seven chapters of text including about 175 typed pages exclusive of tables. Whelpton emphasizes that the need for making estimates was constantly encountered in preparing the various tables from census and registration data available and that the results must be regarded as approximate rather than as exact. For instance, total births are those adjusted to the total United States up to 1933 and for underregistration through-

[7] Whelpton, P. K., "Effect of Increased Birth Rate on Future Population," *American Journal of Public Health*, Vol. 35, No. 4 (April, 1945), p. 330.

[8] Whelpton, P. K., "The Meaning of the 1947 Baby Boom," *Vital Statistics—Special Reports*, Vol. 33, No. 1 (October 7, 1948), p. 5.

[9] National Office of Vital Statistics, "Births by Age of Mother, Race, and Birth Order, United States, 1949," *Vital Statistics—Special Reports*, Vol. 36, No. 9 (October 15, 1951), p. 140.

[10] *Fertility of Marriage*. Census of England and Wales (1911), Vol. XIII (London: His Majesty's Stationery Office, 1917), 477 pp.

out the whole period considered. Estimates were needed, particularly in the early years, for distribution of births by order. They were required at various stages for the preparation of the population bases. For these reasons, work tables are presented in detail and there is much discussion in the text and in the Appendices of the methods used and probable nature of bias entailed.

As explained by Whelpton, "Fertility tables for two types of cohorts . . . 'actual' and 'hypothetical' . . . are presented in this monograph. Each actual cohort consists of white women who were born in the United States during a given 12-month period centering on January 1. For example, the native white women born from July 1, 1899 to June 30, 1900, inclusive, constitute the actual cohort of 1900. . . . In tracing the fertility and mortality experience of each actual cohort through its life the assumption is made [as in the case of life tables] that all members of the cohort were born during the first instant of January 1. This means . . . that . . . the women of the 1900 cohort were exactly 15 years old on January 1, 1915, and exactly 47 on January 1, 1947" (Chapter 2).

For purposes of this study, Whelpton also assumes that the childbearing ages are 15–46 inclusive. Owing to inadequacy of birth registration data before 1920, the actual cohorts are restricted to women within any part of this age group from 1920 to 1949 inclusive. This means that the only actual cohorts that could be considered are those of women born during 1874–1934. Most of the discussion relates to women born during 1890–1930.[11]

The first table gives, for each actual cohort in turn, cumulative birth rates,[12] by order of birth, of women surviving to specified single year of age and calendar year. Extracts of these data are charted in Figures 1 and 2. Figure 1 shows trends from 1920 to the beginning of 1949 in the number of children of given order previously born per 1,000 native-white women attaining ages 25 and 47. It should be noted (a) that each of any twins or other multiple births that occur is given a different order on the birth certificate, so no woman is assigned

[11] Whelpton's monograph also presents cumulative birth rates and distributions by parity of women in "hypothetical" cohorts of each year from 1920 to 1949 but these will not be discussed here. The assumption is made that "the women composing a hypothetical cohort are born on January 1 of the specified year, and have throughout their lifetime the age-specific death rates which all women, and the age-parity specific birth rates which the fecund women who will marry by age 47, actually have in that year" (Table HA). The cumulative birth rates for the hypothetical cohorts presented in Whelpton's tables can be converted into gross reproduction rates (adjusted for parity, sterility, and spinsterhood) by the application of appropriate sex ratios at birth.

[12] Cumulative birth rates "are the numbers of births prior to specified age and date per 1,000 women surviving to that age and date." (Table A.)

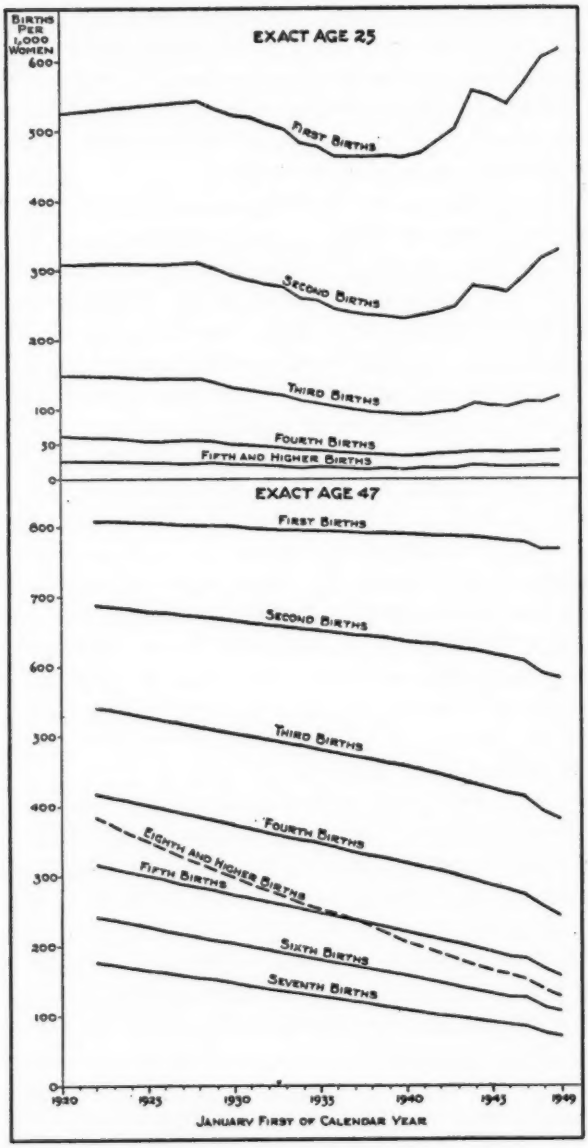

FIGURE 1. Trends since 1920 in numbers of births of specified order occurring prior to ages 25 and 47 per 1,000 native white women surviving to these ages January first of specified calendar year. (Adapted from P. K. Whelpton: *Cohort Fertility: Native White Women in the United States.* Tables for Actual Cohorts. In publication.)

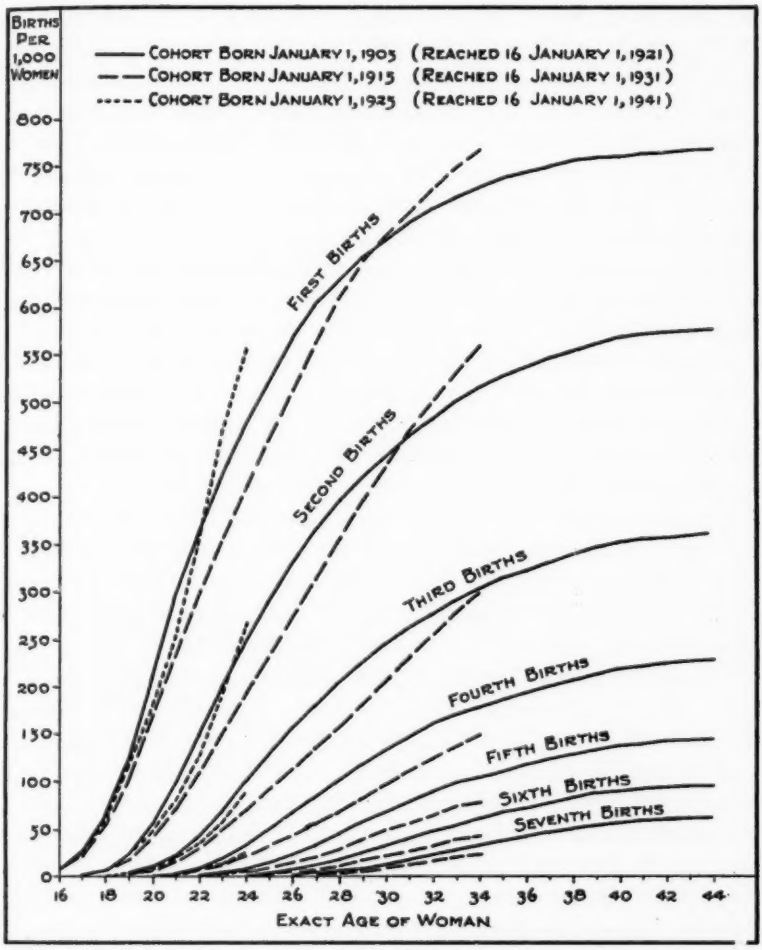

FIGURE 2. Comparison of three actual cohorts of native white women with respect to number of births of specified order occurring prior to given age, per 1,000 women surviving to that age. (Adapted from P. K. Whelpton's data cited under Figure 1.)

more than one birth of a given order; and (b) that women having, say, four births before age 25 are credited not only with fourth births but also with third, second, and first births. Hence the cumulative rates by order of birth can be interpreted *as proportions of women already having at least* one birth, two births, etc. upon reaching any given age.

As indicated in the top section of Figure 1, the number of first births per 1,000 women surviving to age 25 rose slightly from 526 at the beginning of 1920 to 541 at the beginning of 1928, then declined to 462 by 1936 and remained approximately at this level until 1940. Thereafter the trend was upward, until a wartime peak of 557 was reached at the beginning of 1944. Dropping slightly during the next two years the rate climbed again and reached 616 by January 1, 1949 and 625 by January 1, 1950.

Although the actual magnitudes of yearly changes tend to be smaller, the trend in the number of second births per 1,000 women by age 25 follows essentially the same pattern as that described above. The trend in the rate of third births by age 25 also reflects the decline associated with the economic depression and the rise associated with the war. The rate of fourth births by age 25 declined slowly from 62 per 1,000 at the beginning of 1920 to 33 by 1940 and then rose to 39 by 1949 and 40 per 1,000 at the beginning of 1950.

With the arithmetic scale the top section of Figure 1 points up the fact that the absolute declines following 1928 and the absolute increases following 1940 (in cumulative birth rates prior to age 25) become smaller with increasing order of birth. Experimental plotting of the data on semilogarithmic paper indicated that (a) the *rate of decline* following 1928 becomes larger with increasing order of birth, and (b) *rate of increase* following 1940 becomes smaller with increasing order of birth after the first. The 1940–1950 percentage increase in proportion of women having second births by age 25 is higher than that of women having first births by this age.

Although the data for ages between 25 and 47 are not charted, the sharpness of the rise in first births after 1940 decreases with age and virtually disappears by age 40. As indicated in the lower section of Figure 1, there has been a virtually constant downward trend since 1920 in the proportion of women having births of any order by the end of the childbearing period. The relative importance of the decline increases with order of birth and reflects the long-time decrease in size of the completed family. The fertility of women reaching the end of the childbearing period during the past decade, of course, is affected little if at all by the conditions of that period.

Whereas Figure 1 presented cumulative fertility rates for successive cohorts reaching two given ages, Figure 2 follows the survivors of three cohorts through successive ages (beginning with age 16) and shows the number of births of given order (first through seventh) per 1,000 women reaching the successive ages shown on the axis.

The three cohorts chosen for comparison are those of women born at the beginning of 1905, 1915, and 1925 and thus reaching age 16 at the beginning of 1921, 1931, and 1941, respectively. Since the charted data terminate with the beginning of 1949,[13] they relate to ages 16–44 for the 1905 cohort, 16–34 for the 1915 cohort, and 16–24 for the 1925 cohort.

Among women in the 1905 cohort, there were 476 first births per 1,000 women surviving to age 24 (in 1929), 729 per 1,000 reaching age 34 (in 1939), and 770 per 1,000 reaching age 44 (in 1949). The 1915 cohort, reaching age 16 in 1931, belong to the group in which many marriages and births were postponed because of the depression. The rate of first births was consistently lower for this cohort than for the earlier one throughout ages 17–28. Reaching 26 years of age January 1, 1941, these women apparently were still young enough to participate in the marriage and birth boom of that decade. By age 34, reached in 1949, the rate of first births was 767 per 1,000 survivors, higher than that (729) for women of comparable age in the 1905 cohort. In fact, the rate was virtually as high by age 34 for the 1915 cohort as it was by age 44 for the 1905 cohort.

The number of second births per 1,000 women by age 34 was also higher for the 1915 than for the 1905 cohort, but the differences were in the other direction with respect to third and higher orders of birth by age 34.

The fertility of the women born in 1925 is of special interest, since this group reached age 16 in 1941. The occurrence of first births by age 24 was conspicuously high for this group, 553 per 1,000 survivors as compared with 476 for the cohort of 1905 and 409 for the cohort of 1915. The 1925 cohort also outranked the 1905 cohort in rate of second births by age 24, but it fell between the 1905 and 1915 cohorts with respect to third or higher orders of birth by age 24. Thus the 1925 cohort got off to an auspicious start with respect to proportions marrying and having first and second births by age 24. Only time will reveal the fertility behavior of this cohort and other young couples in the future.

The addition of the cumulative birth rates of all orders of birth for a given cohort and age yields the total number of children ever born per 1,000 women surviving to that age. By differencing the cumulative birth rates of successive orders, it is also possible to convert these rates into percentage distributions of women living to a

[13] The data will be extended to the beginning of 1950 in Whelpton's published monograph.

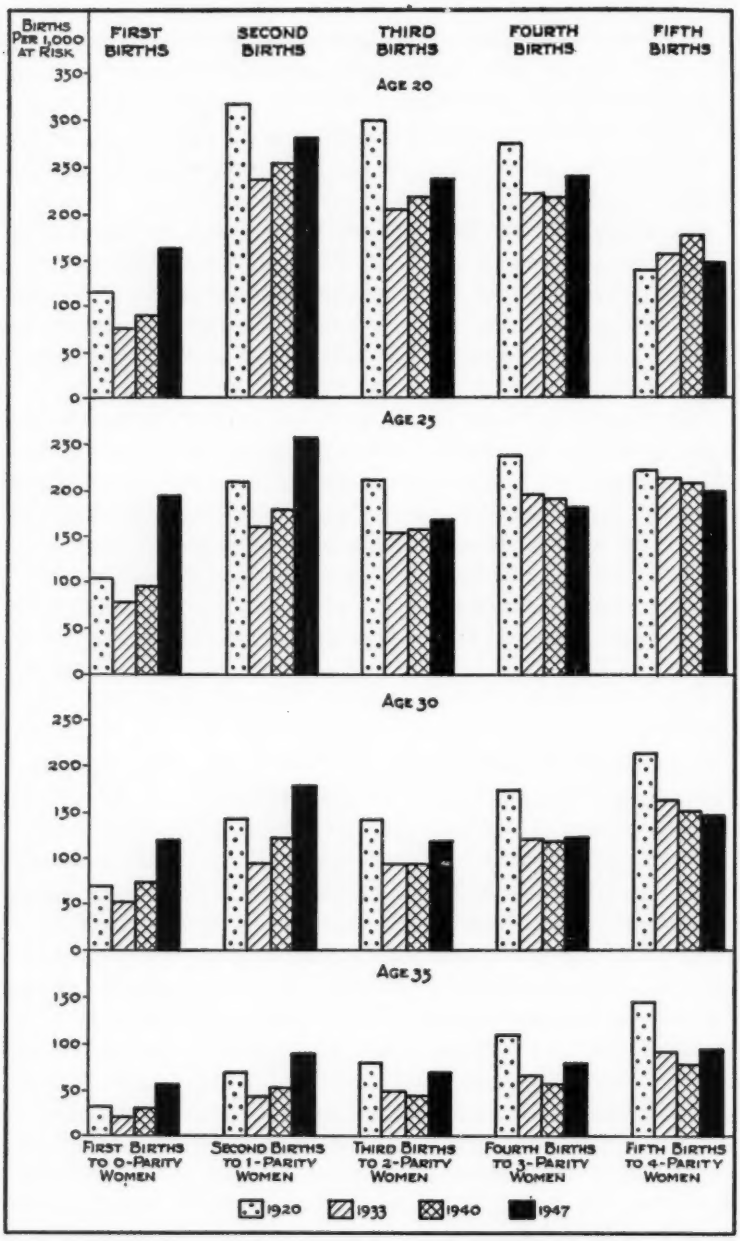

FIGURE 3. Number of births of given order during specified age and calendar year, per 1,000 native white women entering this age and at risk for birth of given order at the beginning of the calendar year. (Adapted from P. K. Whelpton's manuscript cited under Figure 1.)

given age by number of children ever born, by that age.[14] The distributions by parity again point up the heavy role of low-order births in the increase in fertility since 1940. For instance, the total number of births prior to age 30 was 1,453 per 1,000 women in 1941 and 1,786 in 1949. At the earlier date, 35.5 per cent of the women 30 years of age were childless as compared with only 21.8 per cent among women 30 years of age in 1949. The proportion of women with only one child by age 30 was approximately the same (24 per cent) in 1941 and 1949. The proportion with two or three children was 30 per cent in 1941 and 42.5 per cent in 1949. The proportion with four or more children by age 30 was 10 per cent in 1941 and 11 per cent in 1949.

In another series of tables, Whelpton presents annual birth probabilities, by age, by parity of women in actual cohorts. The probabilities are the numbers of births of specified order to women during specified calendar year and age, per 1,000 women at risk for such birth order at the beginning of the year. The rates are computed by relating the number of nth births during the year to the number of women having borne $n-1$ children by the beginning of the year. Figure 3 is based upon data of this type for several specific ages and birth orders for the years 1920, 1933, 1940, and 1947. It will be noted that the probability of first births to zero-parity women (e.g., women who have borne no children) was over twice as high in 1947 as in 1933 for each of the four ages considered. However, the probability of fourth births to three-parity women was actually lower in 1947 than in 1933 among women at age 25, about the same for women at age 30, and somewhat higher for women at age 35. Also, the probability of fifth births to four-parity women at ages 25 and 30 appears to have declined rather continuously since 1933. In most cases, birth probabilities in 1933 were much below those of 1920.

The chapter devoted to the postponing, "making up," and "moving ahead" of births by women in various cohorts is of special interest. In order to estimate the magnitude of birth "deficits" accruing from the depression and of subsequent wartime birth " surpluses," Whelpton first computed numbers of births that would be "expected" with stable economic conditions and no major war after 1930 and then compared the "expected" with the actual numbers. He provides three sets of "expected" births—"high," "medium," and "low."[15]

---

[14] For example, if there are 646 first births and 411 second births per 1,000 women by age 30, there are 235 women per 1,000 (or 23.5 per cent) having one birth but no more by that age.

[15] The "high" assumption was that the average annual percentage decline after 1930 would be half as large as that of 1920–1930 and would stop at the end of 1944. "The 'medium' assumption regarding

The results are striking, especially with reference to first births. As Whelpton indicates, the depression of the 1930's, the prosperity of the 1940's, and World War II caused large changes in the timing of first births to native-white women. The development during 1929–1933 of a "deficit" of between 307,000 and 344,000 first births to women aged 15–39 on January 1, 1934, the increase of the "deficit" to between 519,000 and 609,000 by January 1, 1939, and its virtual elimination by January 1, 1944, point to the large-scale postponement of first births from the first two to the third of these five-year periods.

As for more recent years, Whelpton interprets his findings to mean that the 1947–1950 levels of the crude birth rate are temporary rather than permanent. He points out that about 81 per cent of the changes in first-birth rates may result from changes in marriage rates, since the coefficient of correlation between marriage rates and first-birth rates one year later is about 0.90 for the period under consideration. In view of his estimated "deficits" of fourth and higher births (and possibly of third births) on January 1, 1949 he interprets the rise in numbers of births mainly in terms of an advance in the time schedule of marriage and starting of a family.

Regarding trends during the coming decade, Whelpton points out that to have approximately as many births during 1949–53 and 1954–58 as during 1944–48 will require an unusually favorable combination of circumstances. One group of requirements includes the continuation of prosperity and the avoidance of a major war—in short, the maintenance of conditions which encourage early marriage and the starting of families, and adding of children to families already begun. Another group of requirements includes decreases in the proportions of couples wanting and able to be childless or to have only one or two children which are sufficiently large to balance the increases which seem sure to occur in the proportions able to prevent unwanted fifth and subsequent pregnancies. That the 1950's will meet all of these requirements seems doubtful to him.

It is also important to note, however, that Whelpton's projections, under different assumptions, of the cumulative birth rates by order of birth among cohorts reaching or nearing the end of the childbearing period suggest strongly that the total number of children ever born

---

the 'expected' trend of cumulative birth rates since 1930 is that with stable economic conditions and no major war or other catastrophe the cumulative age-order specific birth rates would have the average annual percentage decrease computed from the difference between the 1920 and 1930 rates." The "low" assumption was that the annual percentage decline after 1930 would be one and one-half times as large as it was during the previous decade.

per native-white woman of completed fertility will reach a low point for the cohort of women born during 1905–1909 and will be somewhat higher for the cohorts of 1910–1914 and 1915–1919. Whether this upward trend will continue among the younger cohorts cannot be predicted with assurance.

Although the readers of Whelpton's monograph doubtless will be impressed by the huge amount of work represented in it, many of them probably will regret that it contains no data for cohorts of married women separately. Whelpton's use of cohorts regardless of marital status has the advantage of yielding fertility trends resulting from the combined factors of change in marriage rates and change in fertility of married women. In the absence of a parallel series of data for married women alone, however, he could do little more than resort to indirect and somewhat crude methods for estimating the relative importance of the two factors. To say, for instance, that about 81 per cent of the changes in first births result from changes in marriage rates because a correlation of 0.90 is found between marriage rates and rates of first births a year later, will not completely satisfy persons schooled in the tradition that correlation does not necessarily imply causation. As already stated, factors stimulating marriages may also stimulate married couples to have a child.

The need for separate data for married women is also keenly felt in the discussion regarding trends in average numbers of births to women of completed fertility. The data relating to all women of completed fertility are highly valuable in their own right. However, it is only when such data relate to married women alone that they can be strictly interpreted as trends in the size of completed family.

The above remarks do not mean that the writer attaches higher value to fertility data relating to cohorts of married women than to cohorts of women regardless of marital status. Both types of data are needed for adequate interpretation.[16]

DIFFERENTIAL FERTILITY RATIOS SINCE 1940

In view of the marked increases in births during the past decade, questions arise concerning the incidence of these increases among different elements of the population and the bearing of this on trends in fertility differentials. It should be emphasized at the outset that the

[16] For a study of changing family size among cohorts of married women in Canada, *see* Ryder, Norman B., "The Cohort Approach: Essays in the Measurement of Temporal Variations in Demographic Behavior," Ph.D. Thesis, Department of Economics and Social Institutions, Princeton University, 1951 (unpublished).

data available for study of these questions are far less satisfactory than those considered in the previous section. For well-controlled analyses of recent trends in fertility differentials, we must await 1950 Census materials of the type which will be described later.

The data presented relate mainly to fertility ratios and replacement rates reported in Census Bureau releases during the past few years and based upon samples of households included in the Current Population Survey. The data have various types of limitations. The samples are too small to maintain through the various analyses adequate breakdowns by such factors as age and rural-urban residence. Furthermore, it must be remembered that since the fertility ratios and replacement indices are based upon numbers of children under 5 years of age, they reflect natality conditions only for the preceding five years. On the basis of the previous discussion there is good reason for believing that the decade of the 'forties was highly abnormal with respect to births. Nevertheless, it is of interest to review the Current Population Survey materials on differential increases in fertility ratios and reproduction rates since 1940.

In general, the increases in fertility ratios since 1940 have tended to be proportionately heaviest among groups previously characterized by lowest fertility and lightest among groups of highest fertility. Thus the percentage increases have been larger in the Northeast than in the South, larger among whites than nonwhites, larger among urban than rural-farm populations, and probably larger in the "upper" than in the "lower" socio-economic classes. By region, the 1940–1947 increases in marital fertility ratios[17] were 26 per cent in the Northeast, 25 per cent in the West, 20 per cent in the North Central States, and only 6 per cent in the South. The increases in the estimated net reproduction rates extended from 41 and 42 per cent for the Northeast and West to 22 per cent for the South.

By color, the fertility ratios increased 18 per cent for whites and 5 per cent for nonwhites from 1940 to 1947. The increases of the net reproduction rates were 33 per cent for the whites and 28 per cent for the nonwhites. In this case the greater proportionate gain in mortality among nonwhites than whites doubtless helps to account for the small white-nonwhite difference in increase of the net reproduction rate.

By type of community, the 1940–1947 increases in fertility ratios

[17] Except those by occupation, the fertility ratios from the Current Population Survey that are discussed in this paper relate to children under 5 per 1,000 *married* women 15–49 years of age. Those by color and residence are not standardized for age. The net reproduction rates and indices of replacement are estimated from children under 5 in relation to all females 15–49 and refer to the five years preceding 1940 and 1947.

were 29 per cent in urban areas, 10 per cent in rural-nonfarm areas, and only 4 per cent in rural-farm areas. The corresponding percentage increases in net reproduction rates were 49, 27, and 12. Woofter[18] has pointed out, however, that higher increases in fertility ratios for urban than for rural areas may not accurately reflect trends in "indigenous fertility" because the ratios for urban areas may be unduly raised by recent influxes of migrants from rural areas. However, one would need to know more about the fertility characteristics of migrants in order to assess the importance of this factor. There is good evidence that migration tends to select the unmarried and some evidence of selective factors in favor of childless couples and small families. The 1950 Census data should provide information on this point.

Preliminary age distributions from the 1950 Census for the country as a whole afford a more comprehensive basis for comparisons of changes in fertility ratios by color and residence than that afforded by samples from the Current Population Survey. It should be emphasized that fertility ratios based upon final figures from the 1950 Census probably will differ somewhat from those in Table 1. It should also be stated that whereas the previous discussion was concerned with changes in number of children under 5 per 1,000 *married* women 15–49 (with husband present), the fertility ratios in Table 1 are based upon children under 5 in relation to women 15–49 *regardless of marital status.* Hence, the percentage changes given in Table 1 are influenced by changes in proportions married to a much larger extent than those previously considered.

As indicated in Table 1, the percentage increase in the general fertility ratio from 1940 to 1950 probably was somewhat higher for the whites than nonwhites in the country as a whole. This, however, appears to be due entirely to differences by color in type of residence. Within each type of residence considered separately, the provisional figures indicate a slightly higher percentage increase in the fertility ratios for nonwhites than for whites. The percentage increase in general fertility ratios, like that of marital fertility ratios, was highest in urban areas and lowest in rural-farm areas. Among the nonwhites, however, the 1940–1950 increase apparently was about the same in rural nonfarm as in rural farm areas.

To return to the Current Population Survey data by education of the woman, the 1940–1947 percentage increase in fertility ratios was highest (34 per cent) for wives reporting four or more years of

[18] Woofter, T. J., Jr., "Trends in Rural and Urban Fertility Rates," *Rural Sociology*, xxiii, No. 1 (March, 1948), pp. 3–9.

college and lowest (12 per cent) for those reporting only 5–6 years in grade school. The slightly higher increase (13 per cent) for the group reporting under five years of grade school was the only exception to a complete direct relation between educational attainment and per-

TABLE 1

NUMBER OF CHILDREN UNDER FIVE PER 1,000 FEMALES 15–49 YEARS OF AGE IN 1940 AND 1950 AND PERCENTAGE INCREASE BY RESIDENCE AND COLOR

Residence and Color	Children Under Five Per 1,000 Females 15–49 Years of Age*		Per Cent Increase 1940–1950 (Preliminary)
	1940	1950 (Preliminary)	
Total U.S.	292	423	44.9
White	287	418	45.6
Nonwhite	339	472	39.2
Urban	227	383	68.7
White	227	382	68.3
Nonwhite	224	394	75.9
Rural Nonfarm	359	492	37.0
White	358	489	36.6
Nonwhite	376	520	38.3
Rural Farm	430	526	22.3
White	409	492	20.3
Nonwhite	535	739	38.1

* The 1940 fertility ratios were computed from age data in U. S. Bureau of the Census: *Sixteenth Census of the United States, 1940. Population*, Vol. II. *Characteristics of the Population* (Washington: Government Printing Office), 1943, pp. 22–25.

The 1950 fertility ratios were computed from preliminary age distributions presented in U. S. Bureau of the Census: *1950 Census of Population, Preliminary Reports. General Characteristics of the Population of the United States, April 1, 1950*, Series PC-7, No. 1 (February 25, 1951), pp. 6–8.

The ratios presented above include no adjustment for under-enumeration of children or for the change in definition of "urban" population in 1950. They are not standardized for age. Final figures for 1950 are expected to differ somewhat from the above since the preliminary age distributions are only approximate (given in thousands only).

centage increase in fertility ratio. A still wider contrast was found with respect to increases in replacement indices which extended from 82 per cent for "college 4+" women to 21 per cent for the "grade school 5–6" group and 29 per cent for the group reporting under five years of school attendance.[19] The much higher increase in the replacement index for "college 4+" women than in the fertility ratio for married women of the same educational level suggests a large increase

[19] U. S. Bureau of the Census: *Current Population Reports—Population Characteristics*, Series P-20, No. 18 (June 30, 1948), pp. 2–15.

from 1940 to 1947 in the proportion of "college 4+" women who are married.

Despite their higher proportionate increase in children, the groups that were characterized by lowest reproduction rates during 1935–1940 tended still to be so characterized during 1942–1947. It is of interest to note, however, that whereas the net reproduction rate for the total Northeast was only 794 per 1,000 (or about 21 per cent below replacement requirements) during 1935–1940, it was 1,123 (or 12 per cent above replacement requirements) during 1942–1947. Similarly, the net reproduction rate for urban areas (726) was about 27 per cent below replacement requirements during 1935–1940 and 8.5 per cent *above* during 1942–1947.[20] The net reproduction rate for the urban groups was about 19 per cent above replacement requirements for the still later period from April, 1944 to April, 1949.

During 1935–1940 the replacement index for women reporting 1–3 years of high school was 997 per 1,000, or almost precisely unity. Above this educational level, reproduction was under replacement requirements. The deficits were 26 per cent for women completing four years of high school, 33 per cent for those attending college 1–3 years, and 48 per cent for these completing four or more years of college attendance. During 1942–1947 the replacement index was below unity only for the "college 4+" group and it was only about 5 per cent below in this instance.[21]

As a result of the trends described above, the magnitude of class differences in fertility ratios has tended to diminish since 1940. An example of this is apparent in Figure 4 showing number of own children under 5 per 1,000 employed men 20–59 years of age, by major occupational group. Except for the relatively high rates for unskilled laborers, the familiar inverse relation of fertility ratios to occupational class of nonagricultural workers is virtually absent. The ratios were lowest for clerical workers in both 1947 and 1949. The fertility ratio for professional men was almost as high as that for semi-skilled workers in 1947 and almost as high as that for skilled craftsmen in 1949. In both years, however, the usual ranking of highest ratios for the farm laborers and next-highest for farmers and farm managers was maintained.

Figure 5 showing fertility ratios to married women 15–49 years of age, by cash income of the family in 1949 and by educational attainment of the woman in 1940 and 1947, yield the familiar inverse relation

[20] *Ibid.*, p. 5.
[21] *Ibid.*, p. 6.

of fertility to socio-economic status. The census release presenting the data by income and occupation contained the following statement: "The findings just discussed [by occupation] may appear to be in conflict to some extent with those shown . . . for women by family income, wherein a consistent inverse relation was found between

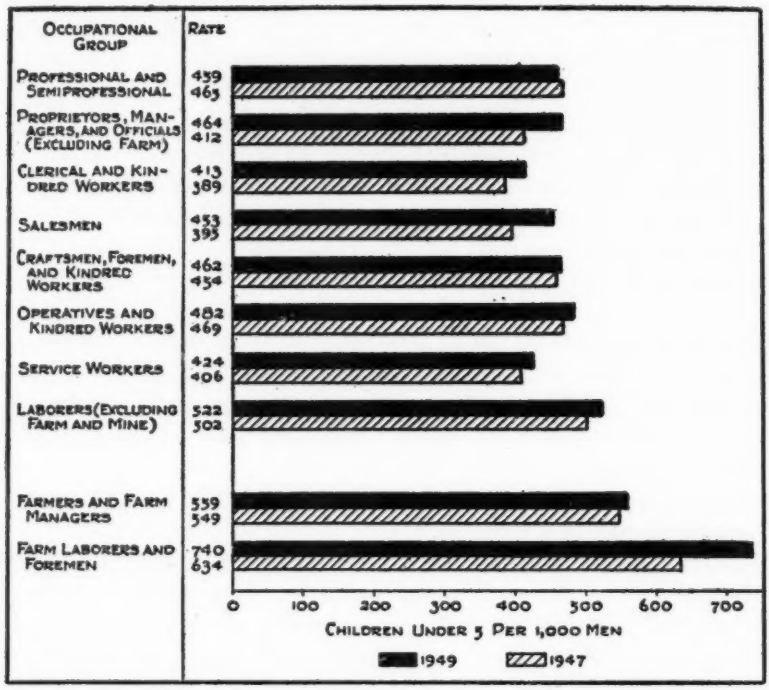

Figure 4. Number of own children under 5 per 1,000 men 20–59 years of age, married and wife present, by major occupational group. Rates standardized for age and relate to civilian employed males in the United States, April 1947 and 1949. (Adapted from U. S. Bureau of the Census: *Current Population Reports—Population Characteristics*, Series P-20, No. 27 (February 3, 1950), p. 11.)

rates of children and economic status. It may be inferred that fertility varies by income within an occupation group in such a manner that when all people of given income are grouped together without regard to occupation, the pattern of inverse association between fertility and income becomes apparent. The failure of the various major occupation groups to show a consistent relation between rates of children and the average income level of a group shows that factors such as the mores and nonpecuniary goals of the group also have an impor-

tant effect on fertility."[22] The last sentence doubtless is true, but a more elementary though perhaps partial explanation might simply be that whereas the classification by occupation rather automatically tends to separate the urban and rural (or at least the agricultural and nonagricultural) workers, the straight classification by income

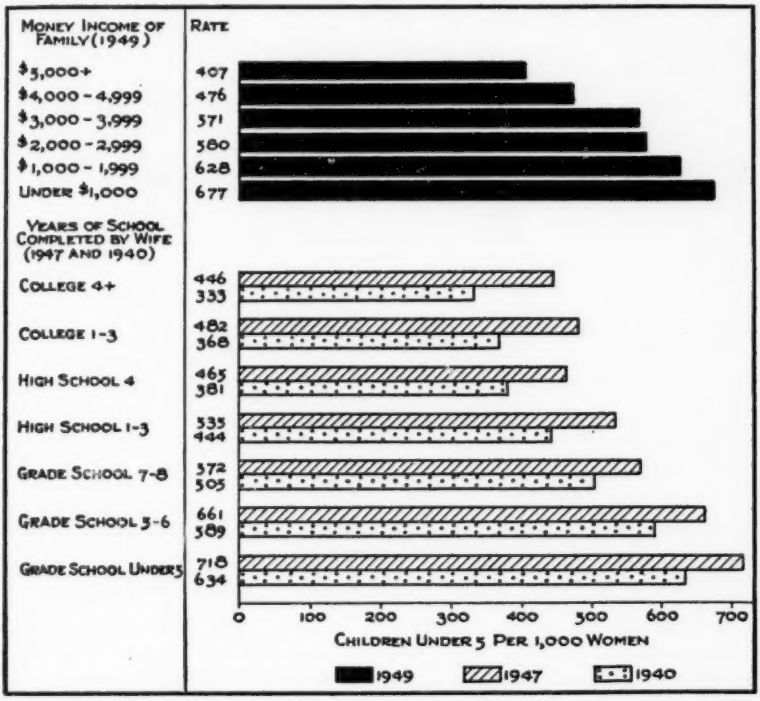

FIGURE 5. Number of children under 5 per 1,000 women 15–49 years of age, married and husband present, by total money income of the family, 1949 (top section), and by years of school completed by the wife, 1947 and 1949 (lower section). Rates standardized for age. (Adapted from U. S. Bureau of the Census: *Current Population Reports—Population Characteristics*, Series P-20, No. 18, June 30, 1948, p. 6 and Series P-20, No. 27 (February 3, 1950), p. 10.)

for the country as a whole does not do this. One might reasonably suppose, particularly since "money income" is used as the basis for classification, that the proportion of rural people in the different income groups would become progressively smaller with increase of

[22] U. S. Bureau of the Census: *Current Population Reports—Population Characteristics*. Series P-20, No. 27 (February 3), 1950, p. 4.

income. Likewise, it would seem safe to assume that the lower educational classes are more heavily weighted with rural people than are the college groups. If this is true, the rural-urban fertility differentials serve to enhance the apparent differentials by income and education in Figure 5.

Limitations of this type are partially removed in Figure 6 showing fertility ratios in 1947 by monthly rent for urban and rural-nonfarm

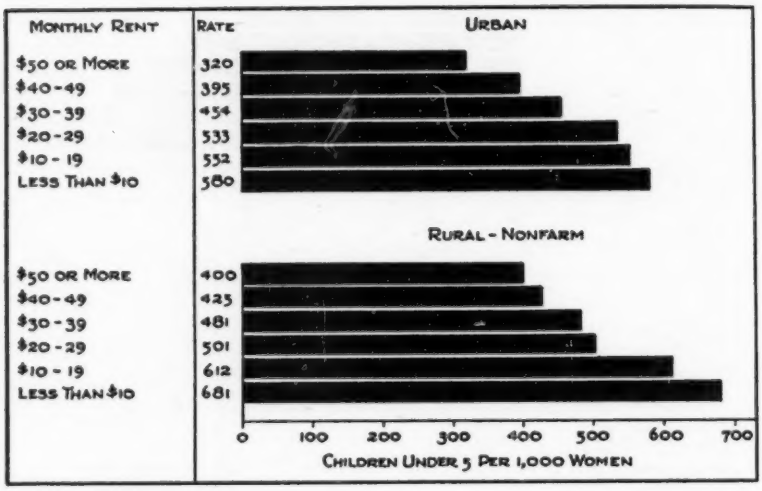

FIGURE 6. Number of children under 5 per 1,000 women 15–49 years of age, married and husband present, by monthly rent and type of residence. Rates standardized for age and relate to renters in urban and rural-nonfarm areas of the United States, April, 1947. (Adapted from U. S. Bureau of the Census: *Current Population Reports—Population Characteristics*, Series P-20, No. 18 (June 30, 1948), p. 17.)

renters separately. For both types of communities, the inverse relation of fertility ratio to monthly rental is clear-cut and consistent. In this instance, however, it seems somewhat unfortunate, particularly for the urban sample, to consolidate rentals of $50 and over. It will be recalled that the 1940 Census data concerning children ever born by monthly rental value of the dwelling unit generally portrayed at least a leveling off, and frequently a reversal, of fertility rates at higher rental-value levels.[23] The decennial census data, of course, permit more detailed

[23] For original data on this point, see U. S. Bureau of the Census: *Population. Differential Fertility 1940 and 1910. Women by Number of Children Ever Born* (Washington: Government Printing Office), 1945, pp. 386–400. For a charting of fertility rates computed from these data and relating to average

classifications than do the relatively small samples covered in the Current Population Survey.

Collectively, however, the materials from the Current Population Survey indicate a sharp contraction since 1940 in various types of group differences in fertility ratios. It is difficult to judge the significance of this in terms of general trends in fertility differentials. As already noted, the small size of the samples prevents sufficient subdivision of the data for adequate analysis. Also, various limitations are inherent in the use of fertility ratios for measures of trends in fertility differentials. The preceding section has indicated that the stream of births during the past decade was greatly affected by the "postponing," "making up" and "moving ahead" of marriages and first births. It may well be that the "postponing" and "making up" of marriages and births were more characteristic of the urban than of the rural and of the "upper" than of the "lower" socio-economic classes. If this is true, the sharp narrowing of differentials in fertility ratios since 1940 may be highly temporary. If cohort data were available, they probably would indicate a considerably less narrowing of differentials since 1940 than that suggested by the fertility ratios.

By way of perspective it is useful to note that whereas any trend toward a contraction of group differences in fertility since 1940 resulted largely from differential *increases* in fertility, in previous years the trend of fertility differentials probably was closely related to the spread of contraception and to the differential *declines* in fertility. If, as seems likely, contraceptive practices originated among the "upper" urban classes and subsequently spread downward in the socio-economic hierarchy and outward to rural areas, the process would yield first a widening and later a narrowing of class differences in fertility.

There is clear evidence of a marked widening of differences in fertility by occupational class in England and Wales during the quarter century preceding 1891.[24] As for later years, Innes' analysis suggests a sharp contraction of differences in birth rates by occupational class from 1921 to 1931.[25] This latter finding is not corroborated by more recent studies made for the Royal Commission on Population but the main report of that body does state that the difference between

---

number of children ever born per ever-married native-white woman 45–49 years of age, by region and type of community, *see* Kiser, Clyde V. and Whelpton, P. K., "Social and Psychological Factors Affecting Fertility. IX. Fertility Planning and Fertility Rates by Socio-Economic Status," The Milbank Memorial Fund *Quarterly*, xxvii, No. 2, April, 1949, p. 189 (Reprint p. 360).

[24] *See* T. H. C. Stevenson's analysis in *Fertility of Marriage*, Census of England and Wales (1911), Vol. XIII, Part II (London, His Majesty's Stationery Office), 1917.

[25] Innes, J. W., "Class Birth Rates in England and Wales, 1921–1931," The Milbank Memorial Fund *Quarterly*, xix, No. 1 (January, 1941), pp. 73–75.

average size of family of the manual and non-manual workers "has perhaps been narrowing."[26]

As for the United States, there are at least suggestions of the enlargement of differences in fertility by occupation during 1850–1910.[27] There are also indications of some diminution of class differences in fertility after about 1925. In his study of 1930 Census materials for the East North Central States on number of children under 5 per couple married 5–9 years, Notestein found that exceptions to the inverse relation of fertility ratios to monthly rental value of the home occurred within the highest rental-value groups.[28] The present writer found analogous exceptions by occupation, education, and family income in his analysis of National Health Survey data relating to births during the preceding year.[29] The more comprehensive 1940 Census data on children ever born indicate similar types of exception to the generally inverse relation of fertility to occupational group, education, and, as previously stated, monthly rental value of the home.[30]

The same type of exceptions to an otherwise inverse relation of fertility of socio-economic status was also found for the total group of "relatively fecund" couples in the Indianapolis Study of Social and Psychological Factors Affecting Fertility. When the analysis was restricted to couples having no pregnancies except those that were deliberately planned by stopping contraception in order to conceive, a rather consistently *direct* relation of fertility to income, occupation, education, and rental value of the home was found. Furthermore, as noted in Figure 7, among these same "number and spacing planned" couples, fertility was *directly* associated with rating of each spouse on

---

[26] Royal Commission on Population, *Report* (London: His Majesty's Stationery Office, 1949), p. 152.

[27] Sallume, Xarifa and Notestein, Frank W., "Trends in the Size of Families Completed Prior to 1910 in Various Social Classes," *American Journal of Sociology*, xxxviii, No. 3 (November, 1932), pp. 398–408.

Kiser, Clyde V., "Trends in the Fertility of Social Classes from 1900–1910," *Human Biology*, v, No. 2 (May, 1933), pp. 256–273.

[28] Notestein, Frank W., "Differential Fertility in the East North Central States," The Milbank Memorial Fund *Quarterly*, xvi, No. 2 (April, 1938), pp. 173–191.

[29] Kiser, Clyde V., *Group Differences in Urban Fertility*. (Baltimore: The Williams and Wilkins Company, 1942.)

[30] *See* footnote 23 above, and U. S. Bureau of the Census: *Population. Differential Fertility, 1940 and 1910. Fertility by Duration of Marriage* (Washington: Government Printing Office), 1947, 338 pp.

Note: Caution must be exercised in comparing 1910 and 1940 Census materials on fertility differentials by occupational class on the basis of data avialable. Both sets of published data are based upon samples and there may be differences arising from changes in urbanization and composition of broad occupational classes. However, among native-white married women, whereas the wives of professional men tended rather uniformly to exhibit lowest age-specific fertility rates throughout ages 20–49 in 1910 in both urban and rural-nonfarm areas, the wives of men of the clerical class frequently exhibited lowest rates in 1940. Also, within urban (but not rural-nonfarm) areas, the proportionate excess in the fertility of the unskilled laborers over that of the professional group was higher in 1910 than in 1940 throughout ages 20–49.

an "index of economic security," which was based mainly upon the self-ratings of each spouse on a variety of specific items such as confidence in meeting future expenses and frequency of facing the possibility of husband's pay-cut or unemployment.

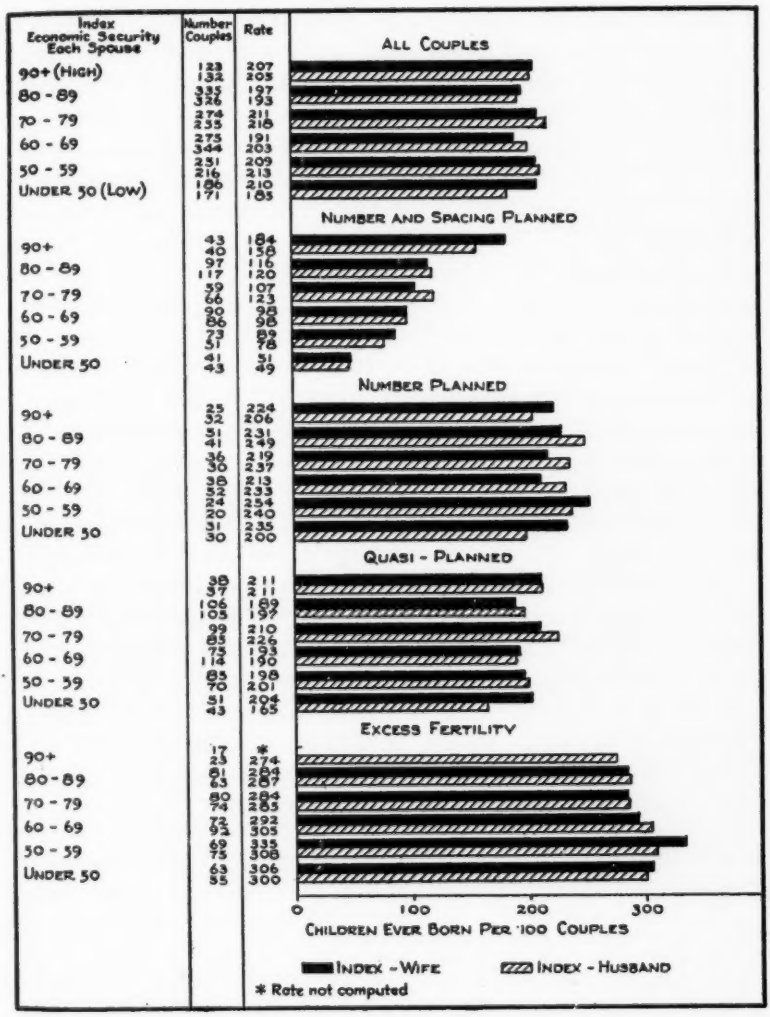

FIGURE 7. Number of children ever born per 100 couples, by fertility-planning status and index of economic security of each spouse. Data relate to 1,444 "relatively fecund" couples in the Indianapolis study. (From The Milbank Memorial Fund *Quarterly*, Vol. xxix, No. 1 (January, 1951), p. 83.)

The 1950 Census secured data on number of children ever born for a 3⅓ per cent sample of ever-married women in the United States. Like the data collected in the 1940 Census, these materials can be analyzed not only by such factors as age, color, nativity, geographic area, type of residence (urban, rural nonfarm, rural farm), but also by measures of socio-economic status, such as occupational class of the husband and educational attainment of the wife, and by refinements regarding times married, age at marriage, and duration of marriage. In addition, data on spacing of children have been prepared for a sub-sample of families in which all children ever born were living at home at the time of the 1950 Census. Although budgetary restrictions may prevent the Bureau of the Census from carrying out tabulations on differential fertility as extensive as those contained in the valuable series of special reports entitled *Differential Fertility, 1940 and 1910*, it is hoped that the Bureau will be able to provide a really substantial body of 1950 data on differential fertility in this country. These data would be of enormous help in interpreting the trends of fertility differentials during the past decade. They would also be of value to persons concerned with other aspects of our population problems.

Although the materials presented in the two sections of this paper provide little basis for judging the future trend of either fertility or fertility differentials, they do suggest that people having fertility under control through either delay of marriage or practice of contraception have contributed more than their proportionate share to the increase in births during the past decade. It seems possible that as the knowledge of means of family limitation becomes more widespread, short-run fluctuations in the birth rate may become more common than they have been in the past. Another implication of the spread of knowledge of fertility control is the increasing importance of learning more about attitudes and motivations regarding family size and how these are affected by current economic and social conditions.

ESTIMATION FOR SUB-SAMPLING DESIGNS EMPLOYING THE COUNTY AS A PRIMARY SAMPLING UNIT*

EMIL H. JEBE†

Iowa State College and North Carolina State College

This paper summarizes a study of the application of various two-stage sampling designs and estimation procedures for preparation of state estimates of agricultural items in North Carolina. Among the principal objectives of the investigation were (i) the comparison of selection of the primary sampling units with equal probability and with probability proportional to size and (ii) assessment of the magnitude of the within primary sampling unit error component contribution to the total survey sampling error. Altogether eight different survey designs with associated estimates have been used to secure information on these two main points. Examination of these designs and estimates indicates a preference for a particular ratio estimate.

INTRODUCTION

THE selection in a two-stage sample survey design of the primary sampling units with probability proportional to size was first presented in 1943 by M. H. Hansen and W. N. Hurwitz. This technique was described in conjunction with a number of other general design principles in their paper, "On the theory of sampling from finite populations" [3]. Since the Hansen and Hurwitz applications related to human populations, this writer's interest was stimulated by a desire to apply this method of selection of the primary sampling units to agricultural surveys and to examine the results which might be obtained under several situations. Recently, the use of the county as a primary sampling unit for agricultural surveys in North Carolina was investigated by L. H. Madow [6]. The work to be reported here also used the county as a primary sampling unit for studying a number of alternative design and estimation procedures. Numerical results reported below are based upon North Carolina data.[1]

* Journal paper No. J-2057 of the Iowa Agricultural Experiment Station, Ames, Iowa, Project 8113, Statistical Laboratory operating with the Bureau of Agricultural Economics, United States Department of Agriculture.

† Presented to the Institute of Mathematical Statistics, December 27, 1950.

[1] The study by Lillian H. Madow [6] and the study reported by the author in this paper comprised two of the projects undertaken by the Institute of Statistics under contract with the Bureau of Agricultural Economics. Even though separately undertaken these two studies were complementary in many respects and similar conclusions could be drawn from numerical results obtained for different purposes. The author's study was not undertaken specifically to evaluate the use of the county as a

For a general description of the population studied reference may be made to [6]. The basic materials used by this study were the same, except for additional sources of data explored by this writer which were:

1. The 1943 N. C. State Farm Census.
2. The BAE Enumerative Surveys [1].
3. The 1945 Sample Census of Agriculture [7].

These sources were added in order to obtain estimates of the "within primary sampling unit" contributions to the error of a sample survey. With respect to stratification, North Carolina was divided into 10 and 20 strata by this writer but equality of size of strata in terms of numbers of farms was not made an aim. This procedure permitted the formation in some cases of somewhat more homogeneous strata than if the strata had been equalized. Sample selection from these more homogeneous strata would tend to force the inclusion of certain types of farming in a sample. Most of the strata used, however, were approximately the same as those shown in [6] (map no. 2, p. 33).

Among the principal objects of investigation in this study were:

1. What effect upon various procedures for sampling an agricultural population will be introduced by selection of the primary sampling units with
   a. Equal probability and
   b. Probability proportional to some measure of size?
2. Of what magnitude are the bias contributions to the "mean square error"[2] for the biased estimates that were chosen for examination?
3. What is the magnitude of the "within county" component of the variance relative to the "between county" component? Will the magnitude of the "within county" component affect the choice of the primary units for different stratifications and methods of estimation?
4. What sample survey designs are to be recommended in terms of the preceding considerations?

DESIGNS AND ASSOCIATED ESTIMATES CONSIDERED

The general type of design being considered is a two-stage scheme in which a *single* large primary sampling unit is selected within each

---

primary sampling unit but rather to examine the probability proportional to size selection technique and the magnitude of within primary sampling unit variance contribution. The county was the most convenient primary unit for which published data were available.

[2] See Hansen and Hurwitz [3], footnote 2, p. 336, for explanation of this concept.

stratum and then a sub-sample is taken within the chosen unit to represent the entire stratum. As already indicated, the county was the primary sampling unit used in this study.

The sub-sample may consist of individual farms or groups of contiguous farms which are taken as clusters or grouped because the headquarters of all farms in the group are defined as being located within a definite area segment.

A few comments on the notation will be necessary in order to make clear the nature of the estimates. Y' is our general notation for an estimate of a population total, Y. To this general notation we add Y' ( ) and insert appropriate letters in the parentheses to identify a particular estimate. The first letter used will indicate the method of selection for the primary sampling units, e.g., e for selection with equal probability and f for selection with probability proportional to number of farms as a measure of size. The second letter used will indicate the method of estimation, i.e., a linear or a ratio estimate. For convenience, omission of the second letter will always indicate a linear estimate. Hence, changes in the second letter will denote a change in the method of forming the ratio for a ratio estimate. An observation for an item on an individual farm is taken as $y_{\alpha ijk}$ with α denoting the stratum, i the county, j the cluster or segment within a county and k the particular farm within a cluster. The range of the subscripts is $\alpha = 1 \cdots H$, $i = 1, \cdots N_\alpha$, $j = 1 \cdots M_{\alpha i}$, $k = 1, \cdots f_{\alpha ij}$. A quadruple summation then yields Y, the population total. It is to be noted that a sample total will involve no summation over i since only a single county is chosen within each stratum. The sample total for a chosen county, and, hence, for the stratum, is indicated as $Sy_\alpha = S_j S_k y_{\alpha ijk}$, the Latin S being used to indicate sample summations. An observation at a previous date on the same characteristic y or a related item is designated as x. Thus, $X_{\alpha i} = \sum_j \sum_k x_{\alpha ijk}$ is a county total and $x = \sum\sum\sum\sum x_{\alpha ijk}$, a population total at a previous date. Sample means are primed, thus, $\bar{y}'_\alpha = Sy_\alpha / Sf_\alpha$ is a county sample mean on a per farm basis while $\bar{\bar{y}}'_\alpha = Sy_\alpha / m_\alpha$ is a county sample mean on a per segment basis. Here Sf_α and m_α are respectively, the number of farms and the number of segments in the sample for a chosen county. When farms are sampled in clusters, the number in a cluster is $f_{\alpha ij}$ and $F = \sum\sum\sum f_{\alpha ij}$ is then the population total of farms. The subscript o is used to indicate a previous total count of farms, thus, $F_{o\alpha}$ for a stratum and $F_{o\alpha i}$ for a county. When counties are to be chosen with probability proportional to a measure of size and number of farms is the measure of size employed, the ratio $F_{o\alpha i} / F_{o\alpha}$ becomes the chance of selection for the ith county in the αth

stratum. An over-all sampling rate or proportion is denoted as t. When the rate is not uniform over the strata, t_α is used.

In several situations we shall wish to indicate the per cent gain in efficiency to be obtained by use of one procedure rather than another. This per cent gain we express as

$$\% \text{ Gain} = \frac{V(a) - V(b)}{V(a)} \times 100$$

in which V indicates a sampling variance and a and b denote the procedures being compared, such as two different methods of estimation or two methods of selection for the primary units. In case the comparison is based on two different stratifications, an appropriate factor, such as 1/2 in going from 10 to 20 strata, must be added as a divisor for $V(a)$ in the calculation. Our approach here follows that given in [6] (note Table 1, p. 37).

We shall now list the principal estimates considered in this study and describe briefly the designs to which they were applied:[3]

1. $Y'(e) = \sum_\alpha F_{o\alpha}\bar{y}'_\alpha = \sum_\alpha Sy_\alpha/t_\alpha$.
   Linear biased estimate.
   p.s.u.—chosen with equal probability.
   sub-s.u.—single farm chosen at random.[4]
   Sampling rate—a uniform rate, $t = \sum_\alpha Sf_\alpha/F_o$ or a separate rate in each stratum, $t_\alpha = Sf_\alpha/F_{o\alpha}$, may be used.
   Expansion factor—estimated total of farms in the stratum (not obtained from the sample, usually a previous total, $F_{o\alpha}$ will be used).
2. $Y'(f) = \sum_\alpha F_{o\alpha}\bar{y}'_\alpha = \sum_\alpha Sy_\alpha/t_\alpha$.
   Linear biased estimate.
   p.s.u.—chosen with probability proportional to estimated number of farms, $F_{o\alpha i}/F_{o\alpha}$.
   sub-s.u.—single farm chosen at random.[4]
   Sampling rate—a uniform rate, $t = \sum_\alpha Sf_\alpha/F_o$, or a separate rate in each stratum, $t_\alpha = Sf_\alpha/F_{o\alpha}$, may be used.
   Expansion factor—estimated total of farms in the stratum (not obtained from the sample, usually a previous total, $F_{o\alpha}$, will be used.
3. $Y'(u) = \sum_\alpha Sy_\alpha/t$.
   Linear unbiased estimate.

[3] For brevity I use here p.s.u. for primary sampling unit and sub-s.u. for sub-sampling unit.

[4] It is assumed that a sufficiently complete list might be available in the chosen county so that such a sample of the ultimate units of observation, farms, could be selected.

p.s.u.—chosen with probability proportional to estimated number of farms, $F_{o\alpha i}/F_{o\alpha}$.

sub-s.u.—an area segment (e.g., Master Sample unit) with farms taken by a headquarters rule.

Sampling rate—controlled in the chosen p.s.u. by the relation, $m_{\alpha i} = tF_{o\alpha}M_{\alpha i}/F_{o\alpha i}$.

Expansion factor—reciprocal of the uniform sampling rate, t.

Here our notation departs from the general description which has been given, $Y'(u)$ is an *unbiased* linear estimate and for this reason the u is used in the parenthesis even though the primary sampling units are selected with probability proportional to estimated number of farms. In fact, this estimate is the linear estimate used in [6], p. 34. The properties of this estimate which make it unbiased are:

a. The utilization of a pre-listing which is available for all counties—the area segments of the Master Sample materials [5], and

b. Control of the sampling rate in terms of segments in the chosen p.s.u. by the relation just indicated above.

4. $Y'(ef) = \sum_\alpha F_{o\alpha} Sy_\alpha / Sf_\alpha$.

Ratio estimate, biased.

Ratio formed—by sample total of item to sample total of farms in the chosen p.s.u.

p.s.u.—chosen with equal probability.

sub-s.u.—an area segment or a cluster of farms. Thus, the Master Sample area segments or any conveniently formable clusters of farms may be utilized.

Sampling rate—$m_\alpha = t_{as} M_\alpha$ where t_{as} is a sampling rate for segments or clusters such that the sample will contain an expected number of farms equal to the desired number of farms for the total sample. Since the sampling rate is not used in the expansion, rigid control is not required as for $Y'(u)$.

Expansion factor—estimated total number of farms in the stratum, usually $F_{o\alpha}$.

5. $Y'(ff) = \sum_\alpha F_{o\alpha} Sy_\alpha / Sf_\alpha$.

The description for this ratio estimate follows exactly that given for $Y'(ef)$, preceding, except that—

p.s.u.—chosen with probability proportional to estimated number of farms, $F_{o\alpha i}/F_{o\alpha}$.

6. $Y'(fx) = \sum_\alpha X_\alpha Sy_\alpha / Sx_\alpha$.

Ratio estimate, biased.

Ratio formed—by sample total of item to a sample total of the same or a closely related item, x, obtained at a previous date for the *identical* segments or clusters.

p.s.u.—chosen with probability proportional to estimated number of farms.

Sampling rate—$m_\alpha = t_{\alpha s} M_\alpha$ (refer $Y'(ef)$, above).

Expansion factor—Stratum total, X_α, of the same or a closely related item, x, which was used in forming the sample ratio.

Here we may note a distinct difference from the ratio estimates $Y'(ef)$ and $Y'(ff)$ described above. For the m_α clusters or segments chosen for the sub-sample in the selected counties, we need for $Y'(fx)$ not only the current data, $Y_{\alpha ij}$, but also the values $X_{\alpha ij}$ of the same or a closely related item for the *identical* segments at a previous date. *These data* are usually not available unless the survey design has been *specifically* set up for this purpose, a *repeat* sampling of the same segments. When this is done the sample ratio, Sy_α/Sx_α, can be formed. The requirement of repeat sampling usually will seriously limit the application of a design employing the estimate, $Y'(fx)$.

7. $Y'(eX) = \sum_\alpha X_\alpha Y'_{\alpha i}/X_{\alpha i}$ or $Y'(eX) = \sum_\alpha X_\alpha M_{\alpha i} Sy_{\alpha i}/(X_{\alpha i} m_{\alpha i})$.[5]

Ratio estimate, biased.

Ratio formed—by sample estimate of p.s.u. total to known p.s.u. total of the same or a related item for a previous date.

p.s.u.—chosen with equal probability.

sub-s.u.—an area segment such as the Master Sample units for which there is a known total, $M_{\alpha i}$, for all counties.

Sampling rate—$m_{\alpha i} = t_{\alpha s} M_\alpha$ where, as for $Y'(ef)$, $t_{\alpha s}$ is an appropriate sampling rate for segments which will yield a desired total sample size in terms of farms.

Expansion factor—stratum total, X_α, for the same or a related item at a previous date.

It should be noted that a ratio estimate of this type, such as $Y'(eX)$, requires no more information from the sample than that required for the linear unbiased estimate, $Y'(u)$, or the ratio to number of farms estimates, $Y'(ef)$ or $Y'(ff)$. Further, in forming the ratio an unbiased estimate is made of the item total for each chosen p.s.u. Thus, this type of ratio eliminates the within p.s.u. bias of a ratio estimate but thereby

[5] It is necessary to retain the subscript i here in order to define clearly the ratio that is formed even though the estimation requires no summation over i just as for the preceding estimates. $X_{\alpha i}$ is a known total for the chosen county and $Y'_{\alpha i} = M_{\alpha i}\bar{y}'_{\alpha i} = M_{\alpha i} Sy_{\alpha i}/m_{\alpha i}$, which makes $Y'_{\alpha i}$ an unbiased estimate of the chosen county total. The capital X is used in the notation $Y'(eX)$ since $X_{\alpha i}$ is used in the estimate instead of Sx_α as for the preceding estimate, $Y'(fx)$.

introduces a larger within p.s.u. variance contribution than a ratio estimate such as $Y'(fx)$, above.

8. $Y'(fX) = \sum_\alpha X_\alpha Y'_{\alpha i}/X_{\alpha i}$ or $Y'(fX) = \sum_\alpha X_\alpha M_{\alpha i} Sy_{\alpha i}/(X_{\alpha i} m_{\alpha i})$. The description for this ratio estimate follows exactly that just given for $Y'(eX)$ except that
   p.s.u.—chosen with probability proportional to estimated number of farms.

The five ratio estimates which have been presented above all use a separate ratio estimate for each stratum instead of a combined ratio estimate from the whole sample of the population. It has been shown that the separate ratio estimate has somewhat greater precision than the combined ratio estimate although the advantage usually is small [2] (p. 125–126).][6] As will be indicated later, there was, in the approach taken by this writer, an additional advantage in the use of a separate ratio estimate for each stratum. Further, it may be observed that ratio estimates $Y'(ff)$ and $Y'(fX)$ are analogous to combined ratio estimates that have been described elsewhere.[7] These ratio estimates in the descriptions generally given appear to be based entirely, although they need not be, upon data obtained from the ultimate sample elements, segments or clusters and farms. Hence, there is an important point to be noted for the ratio estimates $Y'(eX)$ and $Y'(fX)$ presented above. Current sample data are used in forming the numerators of $Y'(eX)$ and $Y'(fX)$ but only primary sampling unit totals are used in the denominators.[8] The diversity of published census materials make readily available such totals for quite similar or closely related items for most sample surveys.

CALCULATIONS FOR EVALUATING THE DESIGNS

After working out the expected values and variances or mean square errors for the above estimates, calculations were undertaken to evaluate

[6] On the other hand, the separate ratio estimate may exhibit a greater bias in some situations due to accumulation of either a positive or negative bias for the individual ratios when summing over the strata. This effect was not noticeable for the situations examined in this study.

[7] e.g., (a) and (b) respectively, given on p. 40 of [6], pp. 118–128 [2] and pp. 13–22 [4]. The lack of comparable data for an estimate such as $Y'(fx)$ was recognised on p. 31 of [6].

[8] This use of previously available information in the denominator of the ratio has inclined me to the use of the term "semi-ratio" for the estimates of this form. Although developed independently, these estimates are not "new"; both E. E. Houseman and R. J. Jessen have considered a similar combined ratio estimate. From another point of view it should also be indicated that this "semi-ratio" estimate attempts to synthesize the observations at a previous date for the segments in the sample in the chosen primary units. Such real data are required for the estimate $Y'(fx)$ described above. The procedure followed with the estimates $Y'(eX)$ and $Y'(fX)$ for synthesizing, in a sense, such data is only one of several techniques that could be suggested for accomplishing this purpose. A similar procedure using a combined regression-estimate was described by Jessen et al. [8].

and compare these estimates. These calculations were finally reduced to relative errors in terms of the coefficient of variation squared, $(CV)^2$, i.e., (Mean Square Error or Variance for Y')/Y^2.

The derivations of these variances, bias terms, and mean square errors will not be given here.[9] In general we desire the expected value of the difference $E(Y'-Y)$ and the difference squared $E(Y'-Y)^2$. The values of these differences for the various estimates are readily obtained by addition and subtraction of appropriate terms. As an example, we show this procedure for $Y'(fX)$.

$$E[Y'(fX) - Y]^2 = E\Bigg[\sum_a \Bigg\{ \left(X_a \frac{Y_{ai}'}{X_{ai}} - X_a \frac{Y_{ai}}{X_{ai}} \right) + \left(X_a \frac{Y_{ai}}{X_{ai}} - E\left(X_a \frac{Y_{ai}}{X_{ai}} \right) \right) + \left(E\left(X_a \frac{Y_{ai}}{X_{ai}} \right) - X_a \frac{Y_a}{X_a} \right) \Bigg\} \Bigg]^2.$$

Carrying out the indicated operations on the three differences within the brackets then readily yields the within county, between county, and bias term contributions to the total error for the estimate, $Y'(fX)$.

With respect to the use of the county as a primary sampling unit essentially similar results for the between county contribution to the sampling error have already been presented in [6]. The principal additional results of calculations which will be summarized below are estimates of the within county contributions to the sampling error for the several estimates and total relative sampling errors including those for the ratio estimates, $Y'(eX)$ and $Y'(fX)$. From a cursory reading of the literature, the impression might be gained that the contribution from the within primary sampling unit component of the variance is small, and, perhaps, unimportant. It is clear, of course, that for any particular stratification and defined sub-sampling procedure the sub-sampling rate can be increased until the between primary unit variance contribution becomes the major or dominating part of the total sampling error. The total sample size required to achieve this situation may become prohibitive in cost in many cases. As will be seen from the numerical results presented below, the North Carolina agricultural population does not exhibit for any reasonably practicable total sample size a within county variance contribution that is relatively small.

[9] These are listed for each estimate in the Appendix following, pp. 68–70. Derivations for each using appropriate notations follow exactly the procedure indicated above for $Y'(fX)$.

Considerable difficulty was experienced in obtaining adequate estimates of this within county contribution to the total survey sampling error. Data were not available for sufficient items to secure a good picture of the within county variability. Obtaining comparability of items for comparison of the several estimates also was a problem.[10] The three sources for the within county data, North Carolina State Farm Census, BAE Enumerative Surveys, and the Bureau of the Census IBM punch cards[11] for Master Sample segments enumerated in the 1945 Sample Census of Agriculture contained only one directly comparable item, farm acres. For one of the estimates, $Y'(fx)$, essentially synthetic estimates of the within county component contributions to the total error were prepared. This was done by borrowing estimated values of ρ from Jessen [4] and using these statistics as estimates of $\rho_{\alpha i}$, the correlation between $Y_{\alpha ij}$ and $X_{\alpha ij}$. These comments indicate the limitations of the within county data that were available for study. It is believed, however, that the information presented is indicative of what might be obtained upon more extensive investigation.[12]

In calculating the between county component contributions to the total error for the ratio estimates, the actual expected values were calculated. This was possible because of the use by this writer of sepa-

---

[10] Some comment on the items selected for inclusion in the study is in order. Obviously, considerations of time, clerical assistance and funds limited the items which could be investigated. Only four items, farm acres, crop acres, value of implements and machinery, and value of land and buildings, are presented for most of the estimates. Quite a number of items were available for analysis in the N. C. State Farm Census data on an individual farm basis (actually, tract basis). It was assumed that the variability pattern for other estimates than $Y'(s)$ and $Y'(f)$ of selected crop items in relation to farm acres and crop acres would be similar to that indicated for $Y'(s)$ and $Y'(f)$. From the available items on the duplicate Census IBM cards, the two value items were selected since it was thought these items might have somewhat different patterns of variability from the crop items over the state of North Carolina. It may be true that these economic value items were inaccurately reported, but it was assumed that this would not seriously affect the variability of the items.

[11] Duplicates of only a part of these cards were available for this study.

[12] It may be of interest to summarize briefly how the estimates of the within county variances were calculated. For estimates $Y'(s)$ and $Y'(f)$ values of s^2 on an individual farm basis were either calculated from data for all farms in a county or estimated by using a large sample (300–500 systematically drawn IBM cards from an alphabetical listing) for a group of 15 widely scattered counties in the State of North Carolina. Results for these 15 counties were deemed sufficiently representative to give a picture of the within county variances for the items studied. For the remaining estimates, $Y'(u)$ through $Y'(fX)$ basic information on the variances within each county was obtained from Bureau of the Census duplicate IBM cards containing area segment summary information for the 1 in 18 systematic sample of the 1945 Sample Census of Agriculture. Sample size from the 1 in 18 sampling rate was adequate in most counties to give stable estimates of the desired variances. On the other hand, use of these data for obtaining estimates of the within county variances may be criticized since these Master Sample segments were systematically selected. Persons intimately acquainted with the Master Sample materials will realize that the method of ordering the segments before the systematic selection was such as to avoid most any systematic pattern in the population in each county. A slightly greater scatter, of course, would be enforced by the systematic procedure, and this would usually increase the calculated variance a bit from that which could be calculated from a random sample, if available. Only estimates of the $\rho_{\alpha i}$, not estimates of the variances themselves, were taken from Jessen [4] for synthesizing estimates of the within county variance contribution for $Y'(fx)$.

rate ratio estimates within each stratum. Where necessary the usual approximate procedure for ratio estimates was used to obtain estimates of the within county component of the error.[13] It is of interest to note that both procedures yielded approximately similar values for the between county sampling variance contributions to the relative errors for $Y'(ff)$ and $Y'(fx)$. In some specific instances, for single strata, considerable differences in value and even sign of the bias term were observed when the bias was calculated by each of the two methods, expected values or the approximate technique.

NUMERICAL RESULTS IN TERMS OF $(CV)^2$

In the following two tables we present estimated relative sampling errors for the designs and estimates which have been described above. The presentation is given first in terms of the coefficient of variation squared, $(CV)^2$ for the total relative sampling error for the item shown. The second table gives the proportion of the total $(CV)^2$ contributed by the within county component of the sampling error. For the biased estimates the bias term is included within the between county component of the error. A third table given later illustrates the magnitude of some of these bias contributions.

The actual estimation situation considered in preparing these estimates of the relative sampling errors was as follows: The 1940 and 1945 U. S. Censuses of Agriculture provided the basic summaries for North Carolina. Then it was assumed that estimates of 1945 state totals for agricultural items were to be made by using the sampling designs and estimates described above. Thus, the 1940 Census data provided the expansion factors, $F_{o\alpha}$ and X_α, and the numbers of farms in the primary units to determine their chances of selection. By combining the 1940 and 1945 agricultural census data estimates of the desired sampling errors and bias components were prepared. Actually, our estimation is only pseudo-estimation—we do not know what the estimate of a particular item total, Y', from a sample survey would have been. Our purpose really is to predict approximately what the sampling error of Y' would be if a sample survey were to be conducted now or in the future for measuring the same item. In doing so, we assume, of course, that variance changes in the population have not been important over the elapsed time period.

An explanation of the sampling rate or sample size in terms of number of farms is needed for understanding the magnitude of the within county error. We have chosen for convenience to apply proportional

[13] Refer [6], p. 40, [2] pp. 118–122 and [9] p. 172 ff. for descriptions of this technique.

allocation uniformly to the strata although we have indicated that the sampling rate need not be equal for all strata for several of the designs. For the 10 strata a total sample of 500 farms was assumed and for 20 strata the number of farms was also doubled, i.e., 1,000 farms comprising the total sample. For the designs employing complete enumeration of Master Sample segments these numbers of farms were translated into 63 and 125 segments, respectively.[14] For comparative purposes we may note that the BAE Enumerative Surveys had for North Carolina an expected number of farms of 731 in 1947 and 512 in 1948.

A review of these tables will permit us to make several comparisons which are of interest for the design of agricultural sample surveys. Of course, these observations are based on North Carolina data and their wider application must take into account the composition of the population to be surveyed. Principally, we shall direct our attention here to the selection of the primary units, use of 10 or 20 strata and total relative sampling error comparisons for the various estimates.

First, we wish to note the over-all gain in efficiency, or decrease in relative error, to be expected when counties are selected with probability proportional to number of farms as the measure of size rather than with equal probability. Our actual measure of size was 1940 number of census farms in each county.[15] There are three situations which we may examine.

1. Estimate $Y'(e)$ vs. $Y'(f)$, Columns 1 and 2, Table 1. Our best comparison here is really on the between county component alone since it was found necessary to use the same within county component for both of these linear estimates. Using the data as they are, we observe a few large gains, some quite small gains and even two cases of losses in efficiency for $Y'(f)$ over $Y'(e)$. Perhaps, a fair estimate of the average total gain for $Y'(f)$ over $Y'(e)$ is about 10 or 15 per cent when we take into account the magnitude of the within county contribution to the total error.[16]

2. The next case which we may look at is the ratio to sample number of farms estimates, $Y'(ef)$ and $Y'(ff)$, Columns 4 and 5, Table 1. Here we are able to make comparisons on both the within county and between county contributions to the total error. There appears to be some gain in efficiency on each component. But again we find that the over-all gains are not large, about 10–15 per cent on the average.

---

[14] The number of census farms per segment in North Carolina is near 8.

[15] Several other measures of size were also examined in this study but they are not included here. Among those studied, number of farms was found to be the best all-round measure of size.

[16] The table given in this paper combines the sampling variance and bias components in the mean square error. Comparisons based on the sampling variance alone give approximately the same results.

3. Our remaining comparison on selection of primary units with equal probability vs. selection with probability proportional to size is that of $Y'(eX)$ with $Y'(fX)$, Columns 7 and 8, Table 1. Here we observe both small gains and small losses in efficiency for $Y'(fX)$ over $Y'(eX)$. Unfortunately, the number of items available for study was

TABLE 1

ESTIMATED RELATIVE SAMPLING ERRORS FOR SELECTED NORTH CAROLINA AGRICULTURAL ITEMS USING EIGHT DIFFERENT SAMPLING DESIGN AND ESTIMATION TECHNIQUES

Item	Design and Estimate							
	1 $Y'(e)$ $(CV)^2$	2 $Y'(f)$ $(CV)^2$	3 $Y'(u)$ $(CV)^2$	4 $Y'(ef)$ $(CV)^2$	5 $Y'(ff)$ $(CV)^2$	6 $Y'(fx)$ $(CV)^2$	7 $Y'(eX)$ $(CV)^2$	8 $Y'(fX)$ $(CV)^2$
North Carolina in Ten Strata								
Farm Acres	.0053	.0052	.0115	.0092	.0088	.0010	.0088	.0087
Crop Acres	.0070	.0067	.0104	.0109	.0095	.0013	.0079	.0072
Cotton Acres	.0648	.0515						
Tobacco Acres	.0351	.0218						
Corn Acres	.0099	.0089						
Value of Implements and Machinery			.0436	.0331	.0284	.0159	.0415	.0387
Value of Land and Buildings	.0107	.0095	.0173	.0150	.0135	.0038	.0120	.0130
Number of Farms			.0066					.0066
North Carolina in Twenty Strata								
Farm Acres	.0023	.0027	.0055	.0043	.0044	.0005	.0045	.0044
Crop Acres	.0039	.0041	.0047	.0055	.0053	.0005	.0037	.0033
Cotton Acres	.0387	.0291						
Tobacco Acres	.0218	.0101						
Corn Acres	.0047	.0041						
Value of Implements and Machinery			.0208	.0154	.0138	.0083	.0233	.0205
Value of Land and Buildings	.0050	.0047	.0079	.0072	.0068	.0018	.0061	.0065
Number of Farms			.0033					.0033

small, but from the available evidence we may conclude that selection of the primary units with probability proportional to size will give only small gains, if any, over selection with equal probability for this type of ratio estimate.

At first glance the results of these comparisons may appear to be quite different from those reported by Hansen and Hurwitz [3]. There are several important points that need to be understood in examining the situations considered by this writer and in [3]. We have already

commented, p. 56, on the part played by the sub-sampling rate within the primary sampling unit in determining the proportion of the total sampling error attributable to the between primary unit variance contribution. Since the major gain, if any, to be expected from the application of the probability proportional to size selection technique is in reduction of the between primary unit variance, it is clear that when the within primary unit variance contribution is relatively large observed gains due to the technique will tend to be small. In this study we have considered only an average sub-sample size of 50 farms in each chosen primary sampling unit (county). This size is not large but increasing it by more than a factor of two would usually be prohibitive in cost for most rural surveys. On the other hand, the size of sub-sample employed for the situation considered by Hansen and Hurwitz was of the order of 300–400 households for most of the primary units. An increase in the sub-sample by a factor of 6–8 in each primary unit would make a considerable difference in the numerical results presented here.

Along with the discussion on the sub-sample size just given we should note that in [3] the results were reported in terms of reduction in the between primary unit variance contribution alone. Larger strata, say of about 30 or more counties each, were employed for the national sample to which the Hansen and Hurwitz principles were applied in [3]. General knowledge of the variance relations for agricultural items would lead one to expect a decreasing relative contribution to the error from the between county variance with decreasing size of stratum. We have used strata whose sizes were of about five or ten counties each. Another important factor in the comparison is that the results reported in [3] relate to unbiased estimates. Among the estimators presented above, only $Y'(u)$ is unbiased. Formation of an unbiased estimator using equal probability of selection for the primary units would result in large variances for the between primary unit contribution in many cases. Should such an unbiased estimator be compared with $Y'(u)$, the gains reported for selection of p.s.u. with probability proportional to size would have been considerably greater. Thus, we may observe that when the above points are taken into account there are no essential differences between the comparisons 1, 2, and 3 discussed above and the comparisons described in [3] on selection with equal probability vs. probability proportional to size for the p.s.u.

For all designs in which the Master Sample segments were used in the the sub-sampling procedure, we find the within county variance contribution to be either about equal to or considerably greater than the between county contribution to the total sampling error (refer Columns

3 through 8, Table 2). This is true even for $Y'(fx)$ which has the lowest within county component of the estimates presented; however, we have noted already that the requirements of this particular design will limit its application. With the general type of design we are considering it appears that the sub-sampling rate would have to be tripled, say a sample of 3000 farms instead of 1000, to reduce the within county

TABLE 2

PROPORTION OF ESTIMATED SAMPLING ERRORS FOR SELECTED NORTH CAROLINA AGRICULTURAL ITEMS DUE TO THE WITHIN PRIMARY SAMPLING UNIT (COUNTY) VARIANCE CONTRIBUTION TO THE TOTAL SAMPLING ERROR

Item	1 $Y'(e)$	2 $Y'(f)$	3 $Y'(u)$	4 $Y'(ef)$	5 $Y'(ff)$	6 $Y'(fx)$	7 $Y'(eX)$	8 $Y'(fX)$
				North Carolina in Ten Strata				
Farm Acres	.396	.404	.739	.652	.636	.500	.932	.943
Crop Acres	.414	.433	.654	.624	.611	.538	.899	.917
Cotton Acres	.125	.157						
Tobacco Acres	.134	.216						
Corn Acres	.283	.315						
Value of Implements and Machinery			.766	.710	.711	.704	.836	.881
Value of Land and Buildings	.346	.389	.711	.533	.578	.632	.867	.892
Nunber of Farms			.864					.864
				North Carolina in Twenty Strata				
Farm Acres	.435	.373	.782	.698	.636	.400	.933	.932
Crop Acres	.385	.366	.681	.564	.509	.600	.919	.939
Cotton Acres	.111	.148						
Tobacco Acres	.092	.198						
Corn Acres	.277	.317						
Value of Implements and Machinery			.788	.753	.710	.687	.815	.868
Valur of Land and Buildings	.480	.404	.785	.569	.574	.667	.885	.908
Number of Farms			.879					.879

variance contribution to a satisfactory level for estimates ($Y'(u)$ and $Y'(ff)$ using 20 strata.

This increase in the sub-sampling rate seems excessive, i.e., beyond what cost considerations would permit for most surveys. If we should choose to change the design as presented there are procedures which will give considerable reduction in the sampling error at a lesser cost than a 3,000 farm sub-sample for estimates $Y'(u)$ or $Y'(ff)$. As an example, we might select two counties per stratum and spread the same size sub-sample (1,000 farm total) over this larger number of primary units. Such a procedure would reduce the between primary unit variance by

a factor of two while maintaining the within primary unit variance at about the same level. Clearly, this would also increase the cost but, perhaps, not in proportion to the increase in primary units. Cost data were not available for assessing the effect of such a change in design. The important fact, however, is that the within county variance component is large for the original Master Sample segments.[17] More homogeneous units need to be defined. Yet, if we look at the within county variances on an individual farm basis (refer Columns 1 and 2 of Tables 1 and 2) we see that it is still large. Classification of large farms into a separate sub-population would help reduce this between farm variation.[18] In reviewing these points and also recognizing that we lack adequate information for examining some possible changes in design it still appears that we need to turn to other estimation methods.

For the ratio estimates $Y'(eX)$ or $Y'(fX)$ an increase in the sub-sample of $1\frac{1}{2}$ to 2 times (1,500 to 2,000 farms) apparently would be sufficient to reduce the within county contribution to the total sampling error to a desired level when 20 strata are used. For $Y'(fx)$ a sub-sampling rate that yields 500 farms probably would be quite adequate for most items. Here we use "desired level" in the sense that the proposed increase in the sub-sampling would reduce the total C.V. to about .05.

With respect to stratification most of the observed gains in going from 10 to 20 strata are small. For $Y'(u)$, Column 3, Table 1, the gains in efficiency range from 8 to 33 per cent for the items shown. Some items presented by Madow [6] (Table 1, column 10, p. 37), exhibited gains of 50 per cent or more.[19] Other estimates presented by this writer, particularly, $Y'(fx)$ and $Y'(eX)$ or $Y'(fX)$, show practically no gains for the increase in number of strata.

Next, we may examine the bias effects for several of the estimates. $Y'(e)$, $Y'(f)$, $Y'(ef)$, and $Y'(ff)$ all depend on number of farms per stratum, $F_{o\alpha}$, as the expansion factor. With increase in the number of strata this source of bias becomes serious and may equal the sampling

[17] It is important to note that the Master Sample area segment information available for this study was based on these area segments as they were used in the 1945 Census of Agriculture. Various techniques can be used to re-define from the Master Sample materials area segments which would be more homogeneous than the units employed for the 1945 Census of Agriculture.

[18] Such a procedure would require developing a separate listing of so-called "large" farms and independent sampling and estimation for this sub-population. From the experience gained with the N. C. State Farm Census data, it appears that quite substantial reductions in variance could be achieved by this separate listing. On the other hand, suitable techniques can also be devised for dealing with the large farms when using an area sample so that reductions in between area segment variability will be obtained.

[19] These items (relating to color of operator, tenure status, and peanuts) show large differences between strata in North Carolina. Hence, larger gains due to increased stratification would be expected. The items used in this study are not concentrated in any part of the state but are fairly widely distributed. With respect to the ratio estimates the observed results also follow expectation in that it has been observed where stratification or ratio estimation are employed, the two techniques will separately achieve reductions in sampling error but when combined the gains are not wholly additive.

variance contribution to the between county component of the error. Table 3 illustrates the magnitude of some of these bias contributions for the various estimates. The other source of bias for the ratio estimates, $Y'(fx)$, $Y'(eX)$, and $Y'(fX)$, expected value of the sample ratio not equal to the true ratio, was found to be practically negligible except in one situation. The item, value of implements and machinery, which changed considerably over the five-year period studied, 1940–1945, gave some bias contributions for $Y'(eX)$ but only negligible amounts for $Y'(fX)$. From this we may observe that the use of selection of primary units with probability proportional to size will aid in reducing the bias for ratio estimates.

TABLE 3

ESTIMATED BIAS CONTRIBUTIONS TO THE MEAN SQUARE ERROR MEASURED AS A PROPORTION OF THE TOTAL MEAN SQUARE ERROR

Estimate	Farm Acres	Crop Acres	Value of Implements and Machinery	Value of Land and Buildings
		North Carolina in Ten Strata		
1. $Y'(e)$	*	.078	†	.156
2. $Y'(f)$	.064	.161	†	.089
4. $Y'(ef)$	*	.05	*	.112
5. $Y'(ff)$	.038	.113	.026	.062
7. $Y'(eX)$	.002	*	.028	.001
8. $Y'(fX)$	*	*	.002	.001
		North Carolina in Twenty Strata		
1. $Y'(e)$	*	.244	†	.249
2. $Y'(f)$	.189	.349	†	.227
4. $Y'(ef)$	*	.174	*	.172
5. $Y'(ff)$	.114	.270	.068	.159
7. $Y'(eX)$	.006	.002	.058	.005
8. $Y'(fX)$	*	*	.008	.001

* Value estimated as less than .001.
† Data not available.

Now, we are ready to make some general comments about the various designs and estimates. First, among the linear estimates we may consider $Y'(u)$ as a standard since it is an unbiased linear estimate. At present it is not feasible to take as a sub-sample a random sample of farms since the *required* lists are not available.[20] If such lists should become generally available, the use of $Y'(f)$ would show a gain of 50 per cent to 100 per cent in efficiency over $Y'(u)$. This is true because of the much smaller within county contribution to the total error for $Y'(f)$. This change in design, however, would increase the cost of enumeration

[20] Limited recent investigations in Iowa have indicated that some county listings may be reasonably satisfactory for certain survey purposes. Such listings can be expected to improve in quality.

for samples of the same total size. All of the designs employing linear estimates would require considerable increases in the sub-sampling rate and perhaps, even further stratification before estimates of satisfactory accuracy (C.V.$\sim$.05) could be obtained. Hence, we must turn to ratio estimates as was indicated by Madow [6], (pp. 38–39).

Turning to the ratio to number of farms estimate, $Y'(ff)$, we find that it offers gains in efficiency of about 15 to 50 per cent over $Y'(u)$. This is helpful but not as much as we would like or need in our situation. Of course, $Y'(fx)$, note Table 1, will about meet our requirements of satisfactory accuracy with 10 strata, or even meet the requirement with a smaller sub-sample when 20 strata are used. But we have already pointed out that a design requiring *repeat* sampling of the identical clusters or segments at two different points in time usually will not be suitable for survey practitioners.

These circumstances led to the study of the ratio estimates $Y'(eX)$ and $Y'(fX)$. These estimates require previous totals of the items to be estimated but only for the chosen county and the stratum; in fact, totals of related items can be used, which makes the estimates quite generally applicable. No repeat samplings of identical elements are required. The gains in efficiency for, say $Y'(fX)$ over $Y'(u)$ are about the same as for $Y'(ff)$. Since the latter may be subject to bias in the expansion factor, number of farms, $Y'(fX)$ would appear to be the preferable estimate. When the number of farms in each stratum at the time of the survey, F_α, is well known, i.e., $F_{o\alpha}$ is a good estimate of F_α, then the estimate $Y'(ff)$ would be preferable to $Y'(fX)$, since $Y'(ff)$ is easier to apply—only one expansion factor is required for all items. Now, let us also look at the effect of increasing the sub-sampling rate. $Y'(fX)$ has a much smaller between county component of the error, thus, doubling the sub-sampling rate will have a greater relative effect for $Y'(fX)$ than for $Y'(ff)$. With a 2,000 farm sample for the state, $Y'(fX)$ might be preferred over $Y'(ff)$ even though the number of farms at the time of the survey is believed to be known accurately.

CONCLUSIONS

The study of these eight designs and associated estimates and other related topics that are relevant to the lay-out of sampling investigations leads us to these observations.

1. The selection of counties with probability proportional to size has not shown large gains in efficiency over selection with equal probability for the items used in this study of the North Carolina agricultural population. This fact does not suggest, however, that the method should not be used. The probability proportional to size selection procedure is not difficult to apply in practice. Even if only small gains in

efficiency can be obtained they are easily secured by a little office work. For linear estimates and ratio to number of farms estimates, which have a large between county contribution to the total error, the gains that can be obtained merit the use of the technique. For the ratio estimate employing previously available county totals, it appears that the gains in reducing the sampling variance may be quite small and that a considerable increase in the sub-sampling rate over the 1 in 287 considered for 20 strata would be required before probability proportional to size selection would exhibit appreciable gains. On the other hand, with these ratio estimates use of the technique appears to be of value for reducing the bias contribution to the mean square error. Thus, even though the within county variance contribution is relatively large, we still find it generally desirable to employ the probability proportional to size selection technique.

2. Bias contribution due to error in the expansion factor may be large relative to the total error. The bias of the ratio estimates $Y'(fx)$ and $Y'(fX)$ was found to be negligible for the situations studied.

3. The within county variance component as a contribution to the total survey error is a factor of major importance for the design of sample surveys of agricultural populations such as that of North Carolina. We have not been able to secure as much information as we would like in terms of number of items and kinds of data for assessing the magnitude of this component. Nevertheless, it appears that considerable attention should be given to the reduction of this contribution to the total sampling error. In particular, the within county variance contribution has been found to be large for the designs requiring area segments in which we employed the 1945 Census of Agriculture Master Sample segments for North Carolina. Efforts should, therefore, be directed toward definition of more uniform area sub-sampling units when estimates $Y'(u)$ or $Y'(fX)$ are to be used.[21]

4. The between county component of the total error for linear estimates with the stratifications employed is so large that we cannot consider linear estimates satisfactory. The ratio to number of farms estimate also exhibits a large between county component of the error although it makes a considerable gain through reduction of the within county contribution to the total error. Hence, we are led to ratio estimates based on designs permitting use of $Y'(fx)$ or $Y'(fX)$. The first of these two we have found undesirable because it requires repeat sampling. Thus, it appears that the ratio estimate $Y'(fX)$ will provide an acceptable estimate for our two-stage sampling design when the within county sampling rate is adequate.

5. Cost considerations have not been introduced into this study.

[21] Refer to footnotes 17 and 18.

Actual survey cost information for two-stage sampling designs needs to be made generally available in order that investigators may have some guides for selection of a particular design and estimate.

These conclusions must be considered in terms of the nature of the population studied and the types of information (data) available for analysis. For the latter, the limitations have been indicated—lack of cost estimates and inadequacy of sources of data for estimating the within county variance contribution. There are alternatives to the conclusions as stated. Unbiased estimates are preferred by some. But even if we increase the number of strata to 40, or used 20 strata with two selections per stratum, and employed a linear unbiased estimate such as $Y'(u)$ we could not attain sufficient reduction in the between county variance component to secure a satisfactory estimate because of the magnitude of the within county variance contribution. Therefore, we have looked for improved methods of estimation even though biased. This same point of view was taken in [6] as noted above. Several ratio estimates have been considered and for the particular North Carolina situations examined, we have found two ratio estimates that exhibited negligible bias. Attention in this study has been directed to the two-stage sampling design employing the selection of a single primary sampling unit at the first stage. It is well known that such a design does not permit unbiased estimation of the survey sampling error from the sample. It has been this writer's experience that most surveys are completed upon preparation of the estimates. Clients of survey agencies are often satisfied with an assurance that the survey error will be "about such and such." Survey researchers are not content with such designs, and, hence, will tend to prefer selection at the primary unit stage of two units per stratum. This may be done when employing selection with probability proportional to an estimated measure of size by selecting two units per stratum with replacement. Such a procedure is appropriate for either a linear or ratio estimate. Simplicity of estimation is also desired by survey designers with a resultant preference for linear estimates or at most a simple ratio to a single item as number of farms of $Y'(ff)$. Yet, more complicated estimates such as $Y'(fX)$ can be rapidly prepared by any statistical clerk as the sample summaries become available. This work can be done in the office and office costs are usually minor with respect to the total costs of a survey.

The author wishes to express his appreciation for the original guidance of Professor W. G. Cochran and the generous assistance of Professor R. L. Anderson in completion of this study. The research reported herein was begun while the author was employed as Agricultural Statistician at the Raleigh Research Laboratory, Bureau of Agricultural Economics, and Resident Collaborator with the Institute of

Statistics, North Carolina State College. A considerable part of the material is taken from the writer's thesis submitted in partial fulfillment of the requirements for the Ph.D. degree in Experimental Statistics, North Carolina State College, October, 1949. Many helpful suggestions and criticisms have been received in the preparation of this paper, particularly from several reviewers and from D. G. Horvitz (now of the University of Pittsburgh) and R. J. Jessen, my colleagues at Iowa State College.

REFERENCES

[1] Brooks, E. M., "The General Enumerative Surveys," *Agricultural Economic Research*. Vol. 1, No. 2: 37–48, 1949.

[2] Cochran, W. G., *Sample Survey Techniques*. Institute of Statistics Mimeo Series No. 7 prepared by the Institute of Statistics, N. C. State College and the Bureau of Agricultural Economics, USDA, Cooperating, Raleigh, N. C., 1948 (revised 1950).

[3] Hansen, M. H. and Hurwitz, W. N., "On the Theory of Sampling from Finite Populations," *Annals of Mathematical Statistics*, 14: 333–362, 1943.

[4] Jessen, R. J., *Statistical Investigation of a Sample Survey for Obtaining Farm Facts*. Research Bulletin 304, Iowa Agricultural Experiment Station, Ames, Iowa, 1942.

[5] King, A. J. and Jessen, R. J., "The Master Sample of Agriculture, I, Development and Use, II, Design." *Journal of the American Statistical Association*, 40: 40–56, 1945.

[6] Madow, Lillian H., "On the Use of the County as a Primary Sampling Unit for State Estimates," *Journal of the American Statistical Association*, 45: 30–47, 1950.

[7] U. S. Bureau of the Census, *Special Report 1945 Sample Census of Agriculture* (Washington, D. C.: U. S. Government Printing Office, 1947).

[8] Jessen, R. J., Blythe, R. H., Jr., Kempthorne, O. and Deming, W. E., "On a Population Sample for Greece," *Journal of the American Statistical Association*, 42: 357–384, 1947.

[9] Deming, W. E., *Some Theory of Sampling* (New York: J. Wiley and Sons, 1950).

APPENDIX

The expressions given below are the $E(Y'-Y)^2$ for each of the designs and estimates presented in this paper (pages 52–55). These expressions give the sampling variances or mean square errors for each estimate. For brevity V is used for this expected value. The notation follows that already indicated on page 51.

1. For $Y'(e)$:

$$V_1 = \sum_{\alpha}\left[F_{o\alpha}{}^2\sum_{i=1}^{N_\alpha}\frac{1}{N_\alpha}\,\frac{F_{\alpha i}-f_{\alpha i}}{F_{\alpha i}}\,\sigma_{\alpha i}{}^2/f_{\alpha i} + F_{o\alpha}{}^2\sum_{i=1}^{N_\alpha}\frac{1}{N_\alpha}\,(\bar{y}_{\alpha i}-\bar{y}_{\alpha s})^2\right]$$

$$+\left[\sum_{\alpha} F_{o\alpha}\bar{y}_{\alpha s} - \sum_{\alpha} F_{\alpha}\bar{y}_{\alpha}\right]^2,$$

where $\sigma_{\alpha i}^2$ = variance for single farms selected at random in the ith county of stratum α and

$$\bar{y}_{\alpha s} = \sum_{i=1}^{N_\alpha} \frac{1}{N_\alpha}\bar{y}_{\alpha i}.$$

2. For $Y'(f)$:

$$V_2 = \sum_{\alpha}\left[F_{o\alpha}^2\sum_{i=1}^{N_\alpha}\frac{F_{o\alpha i}}{F_{o\alpha}}\frac{F_\alpha - f_{\alpha i}}{F_{\alpha i}}\frac{\sigma_{\alpha i}^2}{f_{\alpha i}} + F_{o\alpha}^2\sum_{i=1}^{N_\alpha}\frac{F_{o\alpha i}}{F_{o\alpha}}(\bar{y}_{\alpha i} - \bar{y}_{\alpha f})^2\right] + \left[\sum_{\alpha} F_{o\alpha}\bar{y}_{\alpha f} - \sum_{\alpha} F_\alpha\bar{y}_\alpha\right]^2$$

where $\sigma_{\alpha i}^2$ is defined as for V_1 above and

$$\bar{y}_{\alpha f} = \sum_{i=1}^{N_\alpha}\frac{F_{o\alpha i}}{F_{o\alpha}}\bar{y}_{\alpha i}.$$

3. For $Y'(u)$:

$$V_3 = \sum_{\alpha}\left[F_{o\alpha}^2\sum_{i=1}^{N_\alpha}\frac{F_{o\alpha i}}{F_{o\alpha}}(M_{\alpha i}/F_{o\alpha i})^2\frac{M_{\alpha i} - m_{\alpha i}}{M_{\alpha i}}\sigma_{\alpha i m}^2/m_{\alpha i} + \sum_{i=1}^{N_\alpha} F_{o\alpha}Y_{\alpha i}^2/F_{o\alpha i} - Y_\alpha^2\right],$$

where $\sigma_{\alpha i m}^2$ = variance among clusters (i.e., area segments) selected at random in the ith county.

4. For $Y'(ef)$:

$$V_4 = \sum_{\alpha}\left[F_{o\alpha}^2\sum_{i=1}^{N_\alpha}\frac{1}{N_\alpha}\frac{M_{\alpha i} - m_{\alpha i}}{M_{\alpha i}}V(R_{\alpha i}')/m_{\alpha i} + F_{o\alpha}^2\sum_{i=1}^{N_\alpha}\frac{1}{N_\alpha}(\bar{y}_{\alpha i} - \bar{y}_{\alpha s})^2\right] + \left[\sum_{\alpha} F_{o\alpha}\bar{y}_{\alpha s} - \sum_{\alpha} F_\alpha\bar{y}_\alpha\right]^2$$

where $\bar{y}_{\alpha s}$ is defined as for V_1, above, and $V(R_{\alpha i}')$ = the variance of the sample ratio $Sy_{\alpha i}/Sf_{\alpha i}$ obtained by the approximate procedure (refer footnote 13, above).

5. For $Y'(ff)$:

$$V_5 = \sum_\alpha \left[F_{o\alpha}{}^2 \sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} \frac{M_{\alpha i} - m_{\alpha i}}{M_{\alpha i}} V(R_{\alpha i}')/m_{\alpha i} + F_{o\alpha}{}^2 \sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} (\bar{y}_{\alpha i} - \bar{y}_{\alpha f})^2 \right] + \left[\sum_\alpha F_{o\alpha} \bar{y}_{\alpha f} - \sum_\alpha F_\alpha \bar{y}_\alpha \right]^2$$

where $\bar{y}_\alpha f$ is defined as for V_2 above and $V(R_{\alpha i}')$ is as indicated for V_4.

6. For $Y'(fx)$:

$$V_6 = \sum_\alpha X_\alpha{}^2 \left[\sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} \frac{M_{\alpha i} - m_{\alpha i}}{M_{\alpha i}} V(R_{\alpha iz}')/m_{\alpha i} + \sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} \left(\frac{Y_{\alpha i}}{X_{\alpha i}}\right)^2 - \left(\sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} \frac{Y_{\alpha i}}{X_{\alpha i}} \right)^2 \right] + \left[\sum_\alpha X_\alpha \left(\sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} \frac{Y_{\alpha i}}{X_{\alpha i}} - \frac{Y_\alpha}{X_\alpha} \right) \right]^2$$

where $V(R_{\alpha iz}')$ = the variance of the sample ratio $Sy_{\alpha i}/Sx_{\alpha i}$ obtained as indicated for $V(R_{\alpha i}')$ in V_4, above.

7. For $Y'(eX)$:

$$V_7 = \sum_\alpha X_\alpha{}^2 \left[\sum_{i=1}^{N_\alpha} \frac{1}{N_\alpha} (M_{\alpha i}/X_{\alpha i})^2 \frac{M_{\alpha i} - m_{\alpha i}}{M_{\alpha i}} \sigma_{\alpha im}{}^2/m_{\alpha i} + \sum_{i=1}^{N_\alpha} \frac{1}{N_\alpha} \left(\frac{Y_{\alpha i}}{X_{\alpha i}}\right)^2 - \left(\sum_{i=1}^{N_\alpha} \frac{1}{N_\alpha} \frac{Y_{\alpha i}}{X_{\alpha i}} \right)^2 \right] + \left[\sum_\alpha X_\alpha \left(\sum_{i=1}^{N_\alpha} \frac{1}{N_\alpha} \frac{Y_{\alpha i}}{X_{\alpha i}} - \frac{Y_\alpha}{X_\alpha} \right) \right]^2$$

where $\sigma_{\alpha im}{}^2$ is as indicated for V_3, above.

8. For $Y'(fX)$:

$$V_8 = \sum_\alpha X_\alpha{}^2 \left[\sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} \left(\frac{M_{\alpha i}}{X_{\alpha i}}\right)^2 \frac{M_{\alpha i} - m_{\alpha i}}{M_{\alpha i}} \sigma_{\alpha im}{}^2/m_{\alpha i} + \sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} \left(\frac{Y_{\alpha i}}{X_{\alpha i}}\right)^2 - \left(\sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} \frac{Y_{\alpha i}}{X_{\alpha i}} \right)^2 \right] + \left[\sum_\alpha X_\alpha \left(\sum_{i=1}^{N_\alpha} \frac{F_{o\alpha i}}{F_{o\alpha}} \frac{Y_{\alpha i}}{X_{\alpha i}} - \frac{Y_\alpha}{X_\alpha} \right) \right]^2$$

where $\sigma_{\alpha im}{}^2$ is as indicated for V_3, above.

LATENT STRUCTURE ANALYSIS AND ITS RELATION TO FACTOR ANALYSIS*

BERT F. GREEN, JR.
Massachusetts Institute of Technology†

LATENT structure analysis is a recent contribution to the theory of attitude measurement. It has been formulated and developed by Lazarsfeld [2] as a mathematical model for describing the interrelationships of items in an attitude questionnaire. On the basis of the model it is possible to make some inferences about hypothetical variables assumed to underly the observed responses to the set of items. In its logic as well as in its mathematics, latent structure analysis bears a close resemblance to multiple factor analysis. Before discussing this similarity, however, a brief description of latent structure is in order.

THE LATENT STRUCTURE MODEL FOR DICHOTOMOUS VARIABLES

The latent structure model deals with items having two response alternatives, One of these responses is arbitrarily designated the positive response to the item. If an individual gives the positive response to both items i and j this is referred to as a joint occurrence of positive responses to items i and j, and similarly for any group of items. The manifest, or observable data used in latent structure analysis are the relative frequencies of such joint occurrences:

p_i = proportion of people who respond positively to item i;

p_{ij} = proportion of people who respond positively to both items i and j,

$p_{ij\cdots k}$ = proportion of people who respond positively to all of the set of items $i, j, \cdots, k$.

(These are population values; the sampling problem will not be discussed.)

In the general latent structure model, an underlying attitude distribution $\phi(X)$ is postulated. The model also postulates, for each item, an item conditional distribution, or "trace function" $f_i(X)$. For any X_0, $f_i(X_0)$ is the probability of a positive response to item i, given X_0. It follows immediately that

$$p_i = \int_R f_i(X)\phi(X)dX$$

where R is the region in which X is defined. The fundamental assump-

* Paper presented at the Annual Meeting of the American Statistical Association on December 29, 1950.

† This article was prepared while the author was a Psychometric Fellow at the Educational Testing Service, Princeton, N. J.

tion of latent structure analysis is that the item conditional distributions are independent. For example, the probability of a positive response to both items i and j, given X_0, is $f_i(X_0)f_j(X_0)$. All item interrelationships are due to the hypothetical variable X. The general model may then be represented by the equation

$$(1) \qquad p_{ij\cdots k} = \int_R f_i(X)f_j(X)\cdots f_k(X)\phi(X)dX.$$

Clearly in this very general formulation, very little can be accomplished. Some further assumptions must be made about the trace functions, or the underlying attitude distribution, or both. Present work in latent structure analysis has centered around two particular sets of assumptions, each yielding a fairly flexible model. One assumption in its most general form consists of letting X be a vector in several dimensions, say $X=(x_1, x_2, \cdots, x_m)$, and assuming the $f_i(x)$ to be general polynomials in these x's;

$$(2) \qquad f_i(X) = a_{0i} + a_{1i}x_1 + a_{2i}x_2 + a_{3i}x_1x_2 + a_{4i}x_1^2 + \cdots.$$

This can be obtained from the general model by thinking of the trace functions as expanded in Taylor series.

Two special cases of this formulation are obtained by (1) restricting $f_i(X)$ to a linear function of the x's, and (2) restricting X to a single variable, whereupon the trace function becomes a polynomial in the single variable x. This latter case has been given considerable thought by Lazarsfeld and his associates, [2, 3] with some results being available when the trace functions are linear or quadratic.

In this type of model with polynomial trace functions, the general equation reduces to a relation between the manifest data and some simple function of the coefficients of the polynomials and the moments of the distribution $\phi(X)$. Putting (2) in (1), we have for example,

$$(3) \qquad \begin{aligned} p_{ij} = {} & a_{0i}a_{0j} + (a_{0i}a_{1j} + a_{0j}a_{1i})M_{x_1} + (a_{0i}a_{2j} + a_{2i}a_{0j})M_{x_2} \\ & + a_{1i}a_{1j}M_{x_1^2} + \cdots, \end{aligned}$$

where the M's denote moments of $\phi(X)$.

The conditions under which such equations have unique solutions have not yet been determined, and methods for solving the equations are available only for certain special cases.

A second type of assumption leads to what has been called the latent class model of latent structure analysis. Assume that the region R may be divided into m subregions R_s such that within each subregion the value of each $f_i(X)$ is constant, say $f_i(X)=v_{is}$ for X in R_s. Let

$$n_s = \int_{R_s} \phi(X)dX;$$

that is, n_s is the proportion of the population in the region R_s. Then the general equation becomes

$$p_{ij\cdots k} = \sum_{s=1}^{m} \int_{R_s} f_i(X)f_j(X) \cdots f_k(X)\phi(X)dX$$

$$p_{ij\cdots k} = \sum_{s=1}^{m} v_{is}v_{js} \cdots v_{ks} \int_{R_s} \phi(X)dX \tag{4}$$

$$p_{ij\cdots k} = \sum_{s=1}^{m} n_s v_{is}v_{js} \cdots v_{ks}.$$

In essence, the latent class model assumes that these are m subgroups of people of latent classes with the property that within each class the items are independent, while the manifest item interrelationships are due to differences between the classes. This model is equivalent to assuming a point distribution of attitude, with each point corresponding to a latent class.

Methods of estimating the parameters of this model have been proposed by Lazarsfeld and Dudman [3], Gibson, [3] and Green [1].

GENERALIZED LATENT STRUCTURE—CONTINUOUS VARIABLES

In order to demonstrate the relation of latent structure analysis to factor analysis, it is interesting to consider the generalization of the latent structure model to continuous manifest variables.

The latent class model may be generalized to the case of continuous manifest variables. Let the set of manifest variables be designated by $Y = (y_1, y_2 \cdots y_r)$. Assume that the total population may be divided into m subpopulations, or latent classes, such that within each subpopulation the manifest variables are independent. That is, for class s the frequency function of Y is

$$f_s(Y) = f_{s1}(y_1)f_{s2}(y_2)f_{s3}(y_3) \cdots f_{sr}(y_r).$$

Let

$$\bar{y}_{is} = \int_R y_i f_s(Y)dY = \int_R y_i f_{is}(y_i)dy_i.$$

Then

$$\bar{y}_{is}\bar{y}_{js} \cdots \bar{y}_{ks} = \int_R y_i y_j \cdots y_k f_s(Y)dY.$$

Let the frequency function of Y in the total population be given by

$$g(Y) = \sum_{s=1}^{m} n_s f_s(Y), \quad \text{where} \quad \sum_{s=1}^{m} n_s = 1.$$

Then

$$\int_R y_i y_j \cdots y_k g(Y) dY = \sum_{s=1}^{m} n_s \bar{y}_{is} \bar{y}_{js} \cdots \bar{y}_{ks}. \tag{5}$$

These equations are similar to those of the latent class model for dichotomous variables, with the observed moments of $g(Y)$ taking the place of the observed joint occurrence proportions, $p_{ij\ldots k}$, and $\bar{y}_{is}$ replacing v_{is}. In the continuous case, there are additional equations relating the higher moments of the distributions of the various groups to the corresponding moments of the total group. Thus the matter of the fit of the model is more complex in the continuous case than in the dichotomous case.

The other type of latent structure model which has been discussed also has an interesting analogue in the case of continuous manifest variables. In generalizing the polynomial trace function model we may consider the value of the ith manifest variable, the gth person, say S_{ig}, to be a polynomial in several underlying variables. The resulting equations are identical with those for the case of dichotomous variables, with the joint occurrence proportions being replaced by the moments of the multivariate distribution of the manifest continuous variables; e.g. p_{ij} is replaced by $\sum_g S_{ig} S_{jg}$. One very important case of this analogue is obtained by requiring the S_{ig} to be linear functions of several underlying variables, or factors. That is,

$$S_{ig} = a_{i1} x_{1g} + a_{i2} x_{2g} + a_{i3} x_{3g} + \cdots . \tag{6}$$

This is the basic assumption of multiple factor analysis (Thurstone [4], p. 68).

Thus the generalization of latent structure analysis to continuous data leads directly to multiple factor analysis. This result serves to emplasize the basic similarities of these two models, to which point we may now turn.

THE RELATION OF LATENT STRUCTURE ANALYSIS TO FACTOR ANALYSIS

In the first place, there is a logical similarity in that both models at tempt to account for the observed interrelationships of variables in terms of a few underlying factors, or latent dimensions. In both models it is assumed that all such manifest relationships are due solely to the

mutual relationship of the manifest variables to the underlying factors, latent continua, or latent classes. It is this similarity in basic assumptions which permits equation (6) to represent both the linear factor analysis model and the generalization of a latent structure model.

Although the latent structure model for dichotomous variables uses the concept of the probability of a response, this does not seem to imply any real distinction between the dichotomous and continuous cases. These probabilities merely define the item response distributions, which are analogous with the score distributions (S_{ig}) in the continuous case.

Secondly, due to the common desire to give a simplified description of many variables, a mathematical similarity arises. Both methods attempt to approximate some data matrix by a matrix of low rank. Thus the notion of factoring a matrix is fundamental to both models. In factor analysis it is the correlation matrix which is factored, while in latent structure analysis, it is the matrix of the joint occurrence proportions for pairs of items, $\|p_{ij}\|$. In factor analysis, rank corresponds to the number of factors, or variables needed to account for the correlations; in latent structure analysis, rank corresponds to the number of terms in the trace function polynomials, or to the number of latent classes.

An interesting difference in the mathematics of the two models is that latent structure analysis makes use of the higher-order cross product matrices, such as joint occurrences for triples of items, for quadruples of items, etc. In the latent class case, a unique solution is obtained by using such information. In the polynomial trace function models, some restrictions are placed on the final solution by the higher order data. Factor analysis, on the other hand, does not consider any higher order product moments. It can be shown that under the factor analysts' assumption that manifest variables are *linear* combinations of underlying variables, no additional restrictions are imposed by such higher order data.

The diversity of models in latent structure analysis, implies the possibility of a similar diversity of models in dealing with continuous data. In particular, it emphasizes the possibility of a non-linear factor analysis. Although at present, the problem of non-linear factor analysis has not been solved, it should still be recognized as a possibly better model in some situations. It might be reasonable, in many cases, to assume polynomials of at least the second degree in some of the variables. It is likely, that some of the factor studies which have given poor results might have been clarified if such a non-linear factor model could have been used.

Another model which might be preferable to linear factor analysis for some types of continuous data is the latent class model.

This diversity of models also emphasizes the fact the number of latent dimensions or factors needed to describe the item responses or the test scores is not entirely a matter of mathematics. The number of dimensions used depends upon the assumptions about the functional form of the trace functions, or about the functional relationship of test scores to factors. In linear multiple factor analysis, the approximate rank of the reduced correlation matrix equals the number of factors, or dimensions. If some functions other than linear combinations were admitted, it might be possible to get along with fewer dimensions. In the desire to simplify the explanation of the covariation of a large number of variables, linear functions may not be appropriate in all cases. Methods are needed for deciding when different functions such as higher degree polynomials, or even unpolynomial functions might actually further simplify the results. The decisions as to the type of function to be allowed and the consequent number of dimensions needed are ultimately bound up with the sociological or psychological nature of the data. As such, they are at least partly the responsibility of the social scientist.

SUMMARY

The general model of Lazarsfeld's latent structure analysis has been presented, as well as several more specific models. These models may be generalized to the case of continuous manifest data. In one case, the generalization results in the fundamental equation of linear multiple factor analysis. There is a basic similarity in the fundamental assumptions of latent structure analysis and factor analysis. The two models are also similar in some mathematical respects. The possibility of several models for analyzing continuous as well as dichotomous data is emphasized.

REFERENCES

[1] Green, B. F., Jr., "A General Solution for the Latent Class Model of Structure Analysis," *Psychometrika*, 1951, *16*, 151–166.

[2] Lazarsfeld, P. F., "The Logical and Mathematical Foundation of Latent Structure Analysis," Chap. 10 in: *Studies in Social Psychology in World War II. Vol. IV. Measurement and Prediction*, by Samuel Stouffer *et al.* (Princeton: Princeton University Press, 1950).

[3] Lazarsfeld, P. F. *et al.* Unpublished report of Columbia-Rand project in latent structure analysis. 1950.

[4] Thurstone, L. L., *Multiple Factor Analysis* (Chicago: University of Chicago Press, 1947).

MEASURING MAGNITUDES AND TRENDS IN THE PRODUCTION OF LIVESTOCK AND MEAT*

A. V. NORDQUIST
Bureau of Agricultural Economics

THIS discussion of the subject of measuring magnitudes and trends in the production of livestock and meat will deal with the methods and problems connected with this field in the experience of the United States Department of Agriculture. More specifically, the discussion concerns the activities of the estimating and reporting services of the Bureau of Agricultural Economics.

HISTORY AND DEVELOPMENT

Time will not permit a detailed account of the history and development of the estimating and reporting services relating to livestock and meat. Estimates for the January 1 livestock inventories and value in this country begin with 1867. For more than half a century measurements of annual inventories were about the only statistical activities dealing with livestock on farms. Federal regulations relating to meat inspection resulted in the first continuous series of data on livestock slaughter beginning in 1879. Just after the turn of the present century, interest in livestock statistics was reflected in a number of different kinds of surveys. There was much to learn about production practices relating to livestock. Some of the fundamental relationships of livestock to feed, livestock to meat, supplies and prices were being explored.

Between 1910 and 1920, many mail surveys were made to obtain information on a wide range of subjects relating to livestock. Most of these were of the opinion inquiry kind; that is, informed individuals were asked about changes in inventories, production, marketing weights, losses, disease, and breeding or other management practices. During World War I, vital information was lacking for many phases of the livestock and meat industry. Attempts to satisfy wartime needs were on a patchwork basis. Statistical data on annual changes in inventories did not indicate what was happening in marketing of livestock or in the production of meat. However, some definite ideas as to what was needed had been formed. A fairly complete program for a statistical

* Paper presented at the Annual Meeting of the American Statistical Association, Session: Agricultural Commodity Production and Marketing Statistics, Chicago, Illinois, December 29, 1950.

service on livestock was worked out in cooperation with livestock organizations after funds were provided by Congress for a greatly expanded program. Shortly afterward the pig surveys were started in 1922. They were followed by special reports relating to the lamb crop, wool and mohair production, condition of range feed and livestock, cattle and sheep feeding, and other special phases of the livestock industry. Monthly surveys were also started to obtain data concerning weights and cost of livestock slaughtered under Federal inspection.

One might refer to the decade of the 1920's as the "Golden Age" in livestock statistics. Development of new statistical services on livestock since that time have never approached the achievements of that period. Most of the progress and accomplishments in livestock statistics since the enlarged program began have been in the techniques and procedures used.

For more than 50 years prior to 1920, estimates of inventory numbers were based on opinion or judgment data relating to the changes taking place in inventory numbers. These average yearly reported percentage changes were used to project the most recent census enumerations to estimates for the current year. The concept of obtaining information as to numbers of livestock by asking reporters to give their own individual holdings for the current year and the year preceding was a great forward step. A direct percentage change was calculated from reports on individual holdings and was used in the same way as the old judgment percentage change. Memory or intentional bias was experienced, although it was not fully recognized until later.

The switch from locality judgment to individual farm data as the basic source of estimates emphasized the importance of size, selection, and distribution of samples used. In 1922, arrangements were made with the Post Office Department to distribute questionnaire cards to farmers through the rural mail carriers. The success of this venture in obtaining livestock information from individual farms is portrayed in the rapid development of this survey from a pig survey solely, which it started out to be in 1922, to a survey on all livestock by 1928. A short while later the principal indication of numbers for all States except the Western States was the ratio per farm—the average number of head per livestock farm listed. Matched sampling was also started to get a measure of the percentage change from the previous year. Although the methodology has undergone a marked change, the basic approach of using a census enumeration or a complete count of some kind as a periodic bench mark or starting point and estimating the year-to-year percentage changes to determine the annual estimates is still the same in most of the livestock estimates made by the Department.

KINDS OF SURVEYS AND SOURCES OF DATA

Developments in the methods used for surveying and estimating livestock items have been and still are predominately influenced by time and cost factors. The main objective always has been to make timely estimates. This calls for quick surveys. The width of the field, both for livestock and crops with limited funds made low-cost survey techniques essential. The mailed inquiry thus became the chief means of obtaining survey information. Quick processing of data is also required so that time-consuming enumeration, follow-up, tabulation, and summary methods could not be adopted as current procedures; nor is it possible to make extensive internal analysis of the data at the time the estimates and reports are prepared. Such work becomes research projects on sampling error, sampling techniques, sample distribution, and expansion methods.

The mail survey is not the only method by which information on livestock is obtained. Depending upon the item to be estimated, a variety of methods are used. Interviews with large livestock feeders are necessary in some parts of the country in order to make a reliable estimate of numbers of cattle or sheep on feed. Railroad records of movements into and out of feeding areas are used to determine numbers on feed when the universe is such that mail sampling is difficult and subject to large error. Enumerations of some livestock items such as sows farrowing are available from the State Farm Censuses in a number of important livestock States. In 32 States, records of livestock assessed for taxation are available. These can be used to check past estimates. State brand and health inspection records of livestock moving into and out of the different States are available for about half of the States. Fairly comprehensive and complete records are collected on marketings by State of origin from all important public stockyards and from packers who buy direct from producers. These are among the sources of data that are analyzed and used in connection with preparing or checking the estimates of inventories, production, and marketings of livestock.

It might be well to reflect for a moment on the complexity of the problem of determining how many livestock there are and what changes take place. Livestock numbers change every day of the year. Marketing is a continuous process and this is true also of births and growth of animals. For some species and classes of livestock, disappearance during the year is highly seasonal, as in the case of hogs. Moreover, movements of livestock from one State to another for feeding and pasture purposes are substantial during certain periods. Such movements have no effect on United States totals but are important in the State totals. Thus, if a

measurement of inventories is to be made, it makes considerable difference when the inventory is taken, and what it is to include. To go further, the age of livestock changes as time passes. All this stresses the importance of a constant inventory date if estimates are to reflect the annual trends. It is easy to see that a count of livestock on January 1 would differ materially from a count made on June 1.

As the methods used in estimating the inventories employ a census bench mark as a base or starting point, the census enumeration of livestock is of utmost importance. It would be desirable from the point of view of good livestock statistics to have all census enumerations taken as of January 1. Census enumerations to be good bench marks for livestock estimates need to be complete in coverage and the enumeration taken over a relatively short period of time. When the Census is taken as of April 1, there is the problem of adjusting the census number to January 1 basis before it can be used as a bench mark for estimating the January 1 inventory numbers.

Two methods are used to determine the adjustment factors needed to convert an April 1 Census to a January 1 basis. One method involves a special survey. About the time the Census is taken, a special April 1 inquiry with livestock questions exactly like those of the Census is mailed to respondents who replied to the regular livestock surveys relating to inventories on January 1. These April returns are matched with the January 1 returns. Indicated changes computed from the matched data are applied to the census totals to obtain a "January 1 equivalent number." The other method, adaptable to meat animals, involves using monthly records of livestock movements, marketings and slaughter. These records permit estimates of net disappearance to be made for the period from January 1 to the date of enumeration. The total net disappearance (marketings, slaughter, and allowances for deaths minus imports or inshipments) is added to the census totals to arrive at an indicated number as of January 1.

At present, four principal types of surveys on livestock are made during the year. Foremost among these are the surveys made in cooperation with the Post Office Department through the rural mail carriers in June and December each year. The June survey is the basis for the estimate of the spring pig crop, fall farrowing intentions, the lamb crop, and wool production. The December survey is the basis for estimates of the annual inventory numbers, fall pig crop, and spring farrowing intentions. A mailed survey, called the "Special Inquiry on Disposition of Livestock" is made in January each year and provides information used to estimate the annual calf crop, death losses, and farm

slaughter. Special purpose surveys are made periodically. They deal with inventories of cattle and sheep on feed, the early lamb crop, western sheep numbers and wool production, and reports on condition of range feed, cattle, and sheep. Finally, there are monthly surveys of livestock slaughtered and live weight production in nonfederally inspected plants.

THE RURAL CARRIER SURVEYS

The surveys made through the rural carriers are unique in several

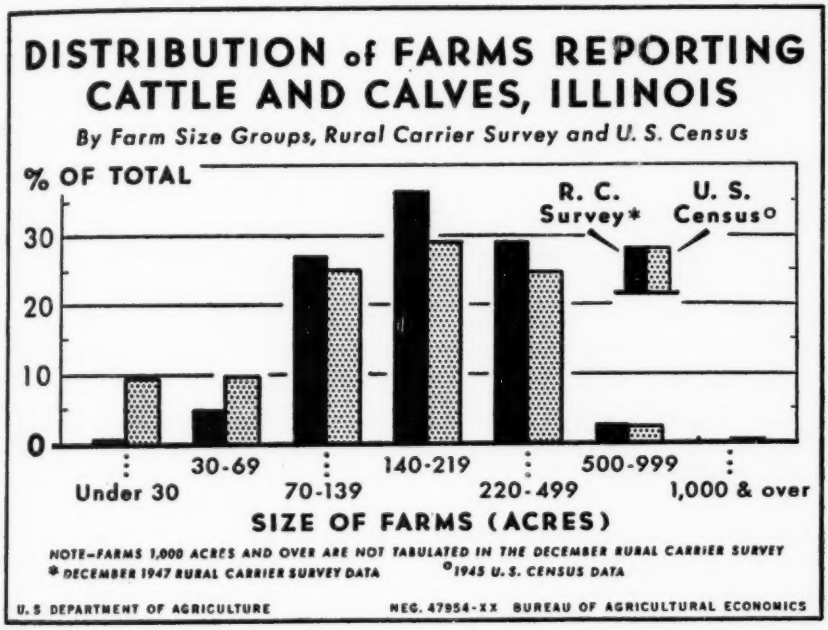

FIG. 1

respects. They are not strictly a mailed inquiry in the sense that the cards are sent to a pre-determined list. Rural carriers are asked to distribute unaddressed cards along their routes so that a cross section of farms is obtained. After the cards are filled out, the carriers collect them and forward them to the State Statistician's office for processing. This sampling procedure is supplemented by direct mailing to obtain more representative samples in areas where the rural mail routes are lacking, sparse, or non-representative. Such supplemental mailing lists are developed from assessment rolls, association membership lists, Government program records, and other sources.

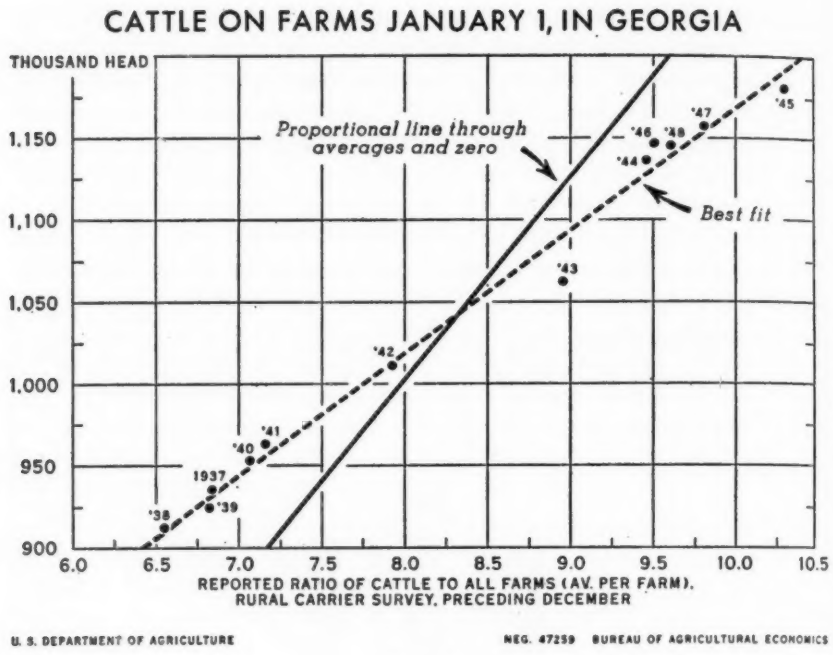
CATTLE ON FARMS JANUARY 1, IN GEORGIA
THOUSAND HEAD
1,150
1,100
1,050
1,000
950
900
Proportional line through averages and zero
Best fit
'45
'47
'46
'48
'44
'43
'42
'41
'40
1937
'39
'38
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.0
10.5
REPORTED RATIO OF CATTLE TO ALL FARMS (AV. PER FARM), RURAL CARRIER SURVEY, PRECEDING DECEMBER
U. S. DEPARTMENT OF AGRICULTURE
NEG. 47259 BUREAU OF AGRICULTURAL ECONOMICS

FIG. 2

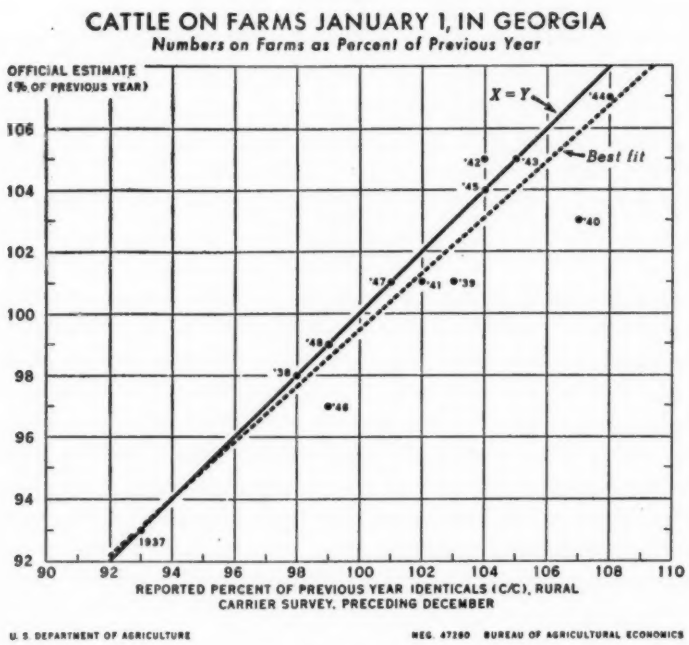
CATTLE ON FARMS JANUARY 1, IN GEORGIA
Numbers on Farms as Percent of Previous Year
OFFICIAL ESTIMATE (% OF PREVIOUS YEAR)
X = Y
Best fit
'44
'42
'43
'45
'40
'47
'41
'39
'48
'38
'46
1937
106
104
102
100
98
96
94
92
90
92
94
96
98
100
102
104
106
108
110
REPORTED PERCENT OF PREVIOUS YEAR IDENTICALS (C/C), RURAL CARRIER SURVEY, PRECEDING DECEMBER
U. S. DEPARTMENT OF AGRICULTURE
NEG. 47260 BUREAU OF AGRICULTURAL ECONOMICS

FIG. 3

The rural carrier distribution of unaddressed cards has the advantage of picking up the new producer or "inner" in the sample. It is possible to match about 40 per cent of the cards with reports from the same producer for the previous year. The distribution of the sample by size of farm resembles the Census distribution, except for the small farm classes (fig. 1). Many of the small farmers are not on rural routes, hence, they do not have an opportunity to fill out a card. This is not serious because only a small number of livestock are on such farms. Moreover, the farms represented in the sample are responsible for most of the annual changes taking place. The selectivity in the sample would be serious if the estimates were determined by an expansion of the average number per farm, using the census number of farms as the universe. However, official estimates are translated from regression charts. By plotting for a series of years the official estimates of actual numbers against the computed average number per farm, an estimate can be read directly from this chart (fig. 2). The estimated number would be determined by the line of best fit. The chart shows a proportional line through average and zero in order to contrast the slope of the line of best fit.

A percentage change in numbers is computed from matched samples of rural carrier survey cards using reports from farmers who reported in both the current and previous year. This indication is also interpreted on a regression chart plotting the annual percentage change in official estimates against the indicated change for a series of years (fig. 3). Figure 3 includes a line, $X = Y$, to illustrate the difference in the slope of the line of best fit and a line, $X = Y$.

In the Western States, the inventory questionnaires are designed to obtain the number of cattle or sheep on hand a year earlier, as well as the current number. This provides an indication of changes between the 2 years. This indication is subject to bias, either memory or intentional. It is possible to measure this bias through the matched sample by comparing totals for the two figures reported for the same year by the same respondent. The percentage change adjusted for bias has been a dependable indication of change in the Western States where it is used.

CATTLE ON FEED SURVEYS AND ESTIMATES

In recent years, considerable effort has been directed toward improving and expanding the statistical and reporting services dealing with cattle on feed. The high-risk aspects of the feeding enterprise require frequent, timely, detailed information on inventories, placements,

and marketings. Special purpose sampling is used in surveying cattle on feed, although in some Western States enumerative surveys and records of livestock movements are the main sources of information. For the Corn Belt States, the questionnaires used have been expanded in recent years to include questions on length of time on feed, and kinds and weight of cattle on feed. Records obtained during the last war on beef production by grades, as well as slaughter of steer and heifer beef by grades, permitted bench-mark estimates to be made for cattle on feed as of January 1, 1945, for Corn Belt States. It was not possible to establish the absolute inventory level for the other dates, April 1 and August 1, on the same basis because production records by grades were not suitable for the periods under consideration and no relationships to total cattle inventories could be studied, as was the case for the January 1 date. Hence, it was necessary to continue to report on the April 1 and August 1 numbers in terms of percentage change rather than absolute numbers.

A research project in three States—Iowa, Illinois, and Nebraska—has made progress in the estimating techniques and statistical services for cattle on feed. It has been possible to measure the actual number of cattle on feed by quarterly dates without benefit of bench mark estimates in the usual sense. Through the use of semi-controlled mail sampling the returns can be directly expanded to an estimate. This estimate can be used later as a projection point for future estimates. So far, procedures have been aimed at independent expansions and estimates for all the quarterly dates.

In semi-controlled mailed sampling, certain data having a high correlation with the item to be estimated are abstracted from the State Farm Census for each prospective respondent who is to receive a cattle-on-feed inquiry. The mailing list is usually a systematic sample drawn from the State Farm Census of individuals who reported on the number of cattle fed, feeder cattle purchased, or fed cattle marketed. With control data taken off at the time the sample is drawn, it is possible to determine in advance how well the sample reflects the universe by comparing the average for the sample as drawn with the universe average for the control item.

Summaries of the returns provide for computing averages for control data, as well as the items to be estimated. Sample indications can then be adjusted by the relationship of the control data for the sample returns to the control data for the universe. The adjusted averages per farm are expanded to totals using the universe number of feeder farms or all livestock farms as the case might be.

One project in Illinois included an interview follow-up on non-response to the mailed inquiry after two mail follow-ups were made. The non-response follow-up survey also was drawn so that it was possible to determine how incomplete was the number of farms reporting cattle placed on feed in the State Farm Census. This additional phase was needed because the feeder question was new on the State census and there was strong evidence of failure to enumerate the item in some townships. Results of the non-response follow-up showed small feeders were not inclined to reply to mail inquiries. An adjustment factor based on the non-response returns was used to adjust the sample averages in lieu of a control data adjustment.

The most recent venture in the Nebraska project, where in and out cattle feeding is important, is the drawing of a supplemental all-farm sample at a lower sampling rate to obtain a measure of the number of new feeders. Control data and repeat mailings are used. The all-farm sample was a systematic drawn sample representing 1.7 per cent of the State Farm Census farms or about 1,500 farms.

In addition to the use of direct sample expansions as indications, matched sampling using the current and previous quarterly returns from the same individuals, as well as those for the current and previous year provide indications of change. Confidence in the methods used has been encouraged by the support given these estimates through check data collected by State of origin on beef steers sold out of first hands for slaughter at the three markets—Chicago, Omaha, and Sioux City. A large part of the fed cattle marketed from these States go through these markets.

MEASUREMENTS OF LIVESTOCK SLAUGHTER AND MEAT PRODUCTION

In measuring magnitudes and trends in the production of meat, the first concern is to determine what might be called the "slaughter universe." The accepted definition for estimating purposes is all establishments and farmers who slaughter meat animals—cattle, calves, hogs, and sheep. The universe breaks down into two distinct parts, (1) farm slaughter and (2) nonfarm slaughter. The latter is called "commercial slaughter."

Farm slaughter takes place on many farms and ranches in the country. The 1945 Census reported 686,000 farms butchering cattle; 412,000 butchering calves; and 3,823,000 butchering hogs—12%, 7% and 65% respectively of U. S. total farms. The procedure used for determining estimates of the number of head slaughtered on farms involves the usual two main steps, (1) establishing a bench mark level

and (2) determining the year-to-year changes taking place. Returns from the January 1 Livestock Disposition Inquiry and other livestock surveys provide indications as to the average number and average live weight of each species slaughtered per farm.

The average slaughter per farm, as computed from the annual surveys, is used to project the trend in slaughter from the Census bench marks. In short, the method employs an enumeration every 5 years and trend charts for translating the annual changes into estimates for intercensal years. Estimates of farm slaughter are made on an annual basis in contrast with "commercial" slaughter which is estimated monthly.

Questionnaires also provide for reporting lard rendered on farms. No information is collected as to the dressed weight of livestock slaughtered on farms. However, data collected during wartime for small slaughterers have been used to estimate the dressing yields for farm slaughter. The information obtained on farm-rendered lard is used in deriving the pork yield for farm slaughter.

The universe for commercial slaughter is composed of two parts, federally inspected and nonfederally inspected plants. The estimates of head and live weight by States represent a combination of enumerated data for part of the universe and estimated totals for the rest of the universe. The Bureau of Animal Industry of the Department of Agriculture, in connection with its regulatory functions relating to meat inspection, compiles a report on the number of animals slaughtered under Federal inspection each month. The number of plants under Federal inspection varies from year to year. The percentage of total commercial slaughter in federally inspected plants varies between months and years also. But as the Bureau of Animal Industry compiles totals for their inspected slaughter, the problem of estimating the total number of head slaughtered commercially centers on the nonfederally inspected plants. At this point, it may be stated that the universe for nonfederally inspected plants is all nonfarm slaughter, except that done in federally inspected plants.

Before World War II, only annual estimates were prepared for nonfederally inspected slaughter. The importance of more frequent reports on this kind of slaughter was forcibly emphasized during the war. Slaughter-control programs included provisions for mandatory reporting by slaughtering establishments. Thus, it was possible to develop a series of monthly data for nonfederally inspected slaughtering plants. Mandatory reporting extended over a period of about $3\frac{1}{2}$ years.

In 1946, Congress provided funds for a program for estimating nonfederally inspected slaughter, so that when OPA controls were removed

and mandatory requirements for reporting no longer existed, reports from slaughterers were obtained on a voluntary basis. The data collected for 1944 and 1945, together with the license lists acquired, were of immense benefit in connection with setting up procedures and estimating techniques to be used in the new program.

In the first place, it was possible to stratify the nonfederally inspected slaughter universe by using data collected during the wartime programs. The stratification by size groups varies by States according to the type of establishment. In general, the universe is broken down into three size groups, wholesale, local, and butcher. An attempt is made to obtain complete coverage of the wholesale group by mailed inquiry, by telephone calls, or by personal visits.

The techniques employed in estimating the number of head slaughtered each month involve separate indications and expansions for each group. A matched sample with current reports matched against reports for the previous month, is used as one of the indications to measure the month-to-month change. Separate estimates for each slaughter group are prepared and estimates for the total are sums of the estimates for the different groups.

Research as to methods of expansion has shown that for both local and wholesale groups a ratio of current slaughter to slaughter during a base period is a more efficient indication for use in expansion to an estimated total than the average per plant. For the butcher group, there is little difference in efficiency between the average per plant and the ratio to base slaughter method of expansion.

To use the ratio to base slaughter method of expansion to an estimate, it is necessary to establish a base for the number of head of each species slaughtered for each plant in the group. Current slaughter is expressed by a percentage of base slaughter for the plants reporting, and this percentage is applied to the total base slaughter of the group to obtain an indication of total slaughter for the group.

Research studies have also been conducted on the length of the base period to be used. It was shown by Hendricks and Hannawald that a base period of 3 months yields almost as high a correlation coefficient as a base period of 12 months.[1]

Indications of the current number of head slaughtered and average live weight for the different species in each group are evaluated and translated to official estimates, using trend charts where bench mark

[1] In service Training No. 33, Bureau of Agricultural Economics Memoranda (Restricted) entitled, "Ratio to Base Method of Making Livestock Slaughter Estimates" prepared by E. B. Hannawald, Division of Livestock and Poultry Statistics, Agricultural Estimates, BAE.

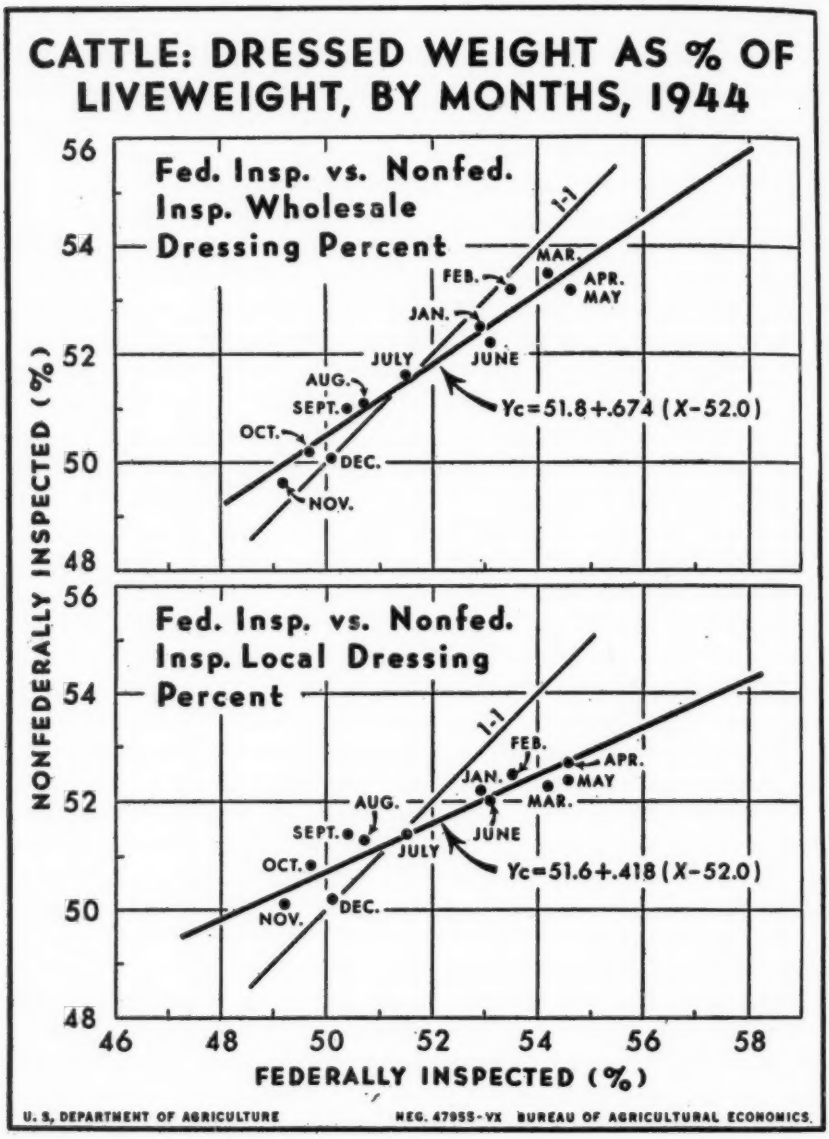

FIG. 4

estimates or other data such as State and municipal records on slaughter are employed to assist in checking on the proper level and month-to-month changes that occur.

The estimates for nonfederally inspected slaughter permit the preparation of estimates of meat production and lard. Information as to live weight and dressed weight of slaughter under Federal inspection has been collected by the Department of Agriculture through voluntary mail returns from packers since 1921. The data as to dressing yields for Federal inspection are used to estimate the dressing yields for nonfederally inspected slaughter as no information is obtained currently

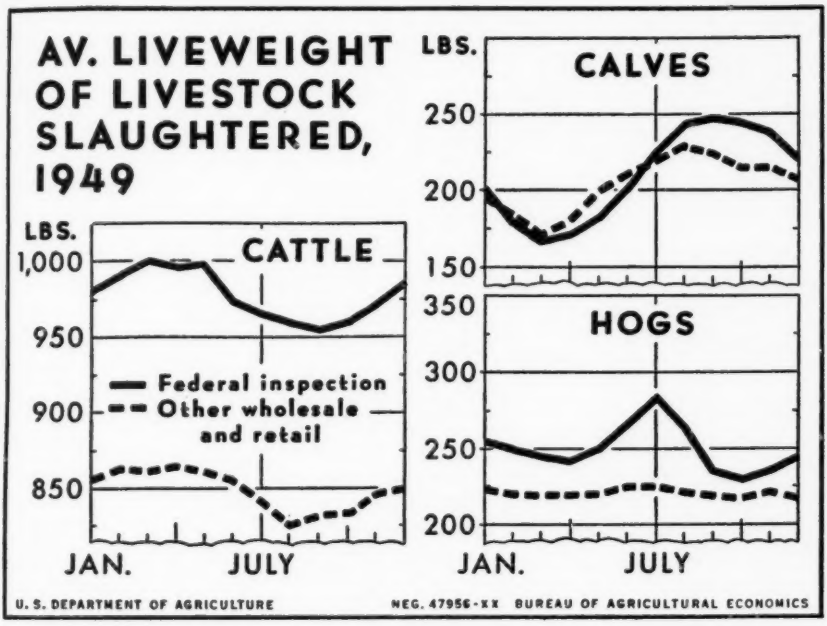

FIG. 5

on dressed weight for the latter group. The current dressing yield for the different species is determined on the basis of the monthly relationship between inspected and noninspected establishments. During the period 1943–45 monthly information on dressed weights and dressing yields was obtained under the slaughter control programs for nonfederally inspected establishments. Comparisons of the dressing yields for inspected and noninspected slaughter revealed that dressing yield for noninspected slaughter tended to follow those for federally inspected establishments, but the changes between months and between seasons were smaller (fig. 4). The same observation is made for trends in average live weight (fig. 5).

SOME THOUGHTS ON FUTURE DEVELOPMENTS

In summary, the "Golden Age" of the 1920's was marked by the development of new statistical series relating to many phases of the livestock industry. In the next two decades, no major additions were made to the existing statistical series and reporting services. During that time, emphasis was placed on sampling and estimating methods. An important field was added to the program when the work on monthly estimates of nonfederally inspected slaughter was started in the 1940's. Also in the last 10 years more detailed data have been developed for the pig crop reports, such as the distribution of sows farrowing by months. Several other reports have included additional data that have proved helpful to users of livestock statistics and reports.

Research in the field of mail surveys and estimating methods has introduced more effective techniques and procedures. Studies continue on factors affecting response, questionnaire design, and sampling and estimating methods. Experimentation in all these fields needs to be continued so that improvements can be made. Periodic surveys and studies need to be made to verify the relationships of dressing yields used in determining production of meat.

The demand for statistics on livestock and meat is far from satisfied. No forecast is possible as to the rate of progress along this line, but the direction of future developments is indicated by the need for certain information and reports.

A reflection of the future may be found in the Federal and State Research and Marketing Act projects being conducted in several States. Programs in individual States include current estimates of sows farrowing by months, data relating to marketings of livestock, breeds of livestock, and estimates of inventories by counties. Other types of information proposed or being considered under a State program for study and development include monthly intentions to market and more current information on the calf crop.

Judging from the past, as the new projects become permanent additions to the over-all program, even further requirements will come to the front in this continuous evolution.

REPRINTS OF ABSTRACTS IN STATISTICAL METHODOLOGY

Edited by MAX A. WOODBURY, Princeton University

POPULATION INDEX 1951

1819 Adamo, M.

Statistical contribution to the study of sudden deaths in Italy. (Contributo statistico allo studio della morte improvvisa in Italia.) Atti dell'Accademia dei Fisiocritici di Siena, Sezione Medico-Fisia 18(2): 285–298. 1950.°

1980 Arrow, Kenneth J., and Arrow, Selma S.

On Mandelbaum's study of "The industrialization of backward areas." U.S. Air Force, Project Rand, Research Memorandum RM-373. Santa Monica, Rand Corp., 1951. 20 pp.

A study of the basis afforded by Mandelbaum's logical structure and presuppositions for quantitative prediction and planning in backward areas, including the determinants of the flow of labor and the assumptions that have to be made to estimate the coefficients needed.

2178 Hosemann, H.

The bases of statistical methods for medical students and biologists. (Die Grundlagen der statistischen Methoden für Mediziner und Biologen.) Stuttgart, Georg Thieme, 1949.°

PSYCHOLOGICAL ABSTRACTS 1951

5001. **Iyer, P. V. Krishna. A non-parametric method of testing k samples.** *Nature, Lond.*, 1951, **167**, 33.—"The object of the present communication is to show that S of Kendall's r for rank correlation can be used in testing the null-hypothesis for k samples from univariate populations."—*A. C. Hoffman.*

5829. **Birnbaum, Z. W.** (*U. Washington, Seattle.*) **On the effect of the cutting score when selection is performed against a dichotomized criterion.** *Psychometrika*, 1950 **15**, 385–389.—The method which is discussed and illustrated, applies to such a selective process as that involved when an admission test is used to select candidates for training and, after completion of the course, the subjects are judged successful or unsuccessful.—*M. O. Wilson.*

5834. **Carroll, John B., and Bennett, C. C.** (*Harvard U., Cambridge, Mass.*) **Machine short-cuts in the computation of chi-square and the contingency coefficient.** *Psychometrika*, 1950, **15**, 441–447.—Rapid routines for computing chi-square tests of the following are presented: (1) goodness of fit between an observed and a theoretical distribution, (2) independence of several distributions displayed in an $r \times c$ table, and (3) independence of distribution in an $r \times c$ table.—*M. O. Wilson.*

5846. **Jenkins, William Leroy.** (*Lehigh U., Bethlehem, Pa.*) **Quick estimation of multiple R.** *Educ. psychol. Measmt.*, 1950, **10**, 346–348.—A method and an example of a short-cut method for estimating the multiple R for a test battery. It is felt that "even if a Doolittle solution is finally obtained, the method provides a preliminary estimate and a useful cross-check against serious blunders in computation."—*J. E. Horrocks.*

5857. **Votaw, D. F., Jr., Rafferty, J. A., and Deemer, W. L.** (*USAF School of Aviation Medicine, Randolph Field, Tex.*) **Estimation of parameters in a truncated trivariate normal distribution.** *Psychometrika*, 1950, **15**, 339–347.—Maximum-likelihood estimators for certain parameters when values of others are known are featured. Approximate variances and covariances of the estimators, when the sample is large, are given.—*M. O. Wilson.*

7176. **Mosteller, Frederick.** (*Harvard U., Cambridge, Mass.*) **Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations.** *Psychometrika*, 1951, **16**, 3–9.—Thurstone's Case V of the method of paired comparisons (see 1: 2504), assumes among other things zero correlations between sensations resulting from pairs of stimulation. It is shown in this study that this assumption may be relaxed to that of equal correlations between these pairs without changing the method. 4 references.—*M. O. Wilson.*

BOOK REVIEWS

Statistical Methodology Reviews, 1941–1950. *Edited by Oscar Krisen Buros* (Director, Institute of Mental Measurements, Rutgers University). New York: John Wiley and Sons, Inc., 1951. Pp. xii, 457. $7.00.

ESSENTIALLY, this volume is a collection of reviews of books on statistics published in the last ten years. According to the Preface, "it lists ... three hundred and forty-two statistical methodology books written in English and published or reviewed in the ... period 1941–1950 in ... Australia, Canada, Denmark, England, Holland, India, Puerto Rico, Scotland, South Africa, Sweden, Switzerland, and the United States ... [It] includes ... eight hundred and forty-two review excerpts from one hundred and twelve journals."

"This volume differs in several respects from earlier volumes in the series" of which it is the third. It "lists only books of statistical methods and such closely related subjects as probability and mathematics of statistics. ... The reviews that seemed least valuable have been omitted ... [and] some important reviews have been omitted in order not to give excessive space to any one book. ... The name index has been changed to permit rapid identification of authors and reviewers. 'The Classified Index to Books' has been extended to include a listing of authors and publication dates for each subject."

The format and typography are excellent, and the price seems modest for so large a volume with so small a potential market—almost exclusively libraries.

In preparing this notice an instance was noted where, contrary to Buros' usual criteria in selecting excerpts, only the descriptive part of a review was quoted, and an important though small criticism was omitted. This was Margaret A. Labadie's notice of *Tables of the Binomial Probability Distribution*, issued by the National Bureau of Standards, in this JOURNAL (Vol. 45, 1950, p. 474). Mrs. Labadie mentioned an error in one of the tabular entries, which is not quoted by Buros.

Another minor error was noticed in entry number 280, the Statistical Research Group's *Sequential Analysis of Statistical Data: Applications*, where an additional 47-page Section 7 is referred to as "an additional 47-page supplement to Section 7."

W. A. W.

An Introduction to Statistical Analysis. *Wilfrid J. Dixon* (Associate Professor of Mathematics, University of Oregon) and *Frank J. Massey, Jr.* (Assistant Professor of Mathematics, University of Oregon). New York: McGraw-Hill Book Company, Inc., 1951. Pp. x, 370. $4.50.

WILLIAM G. COCHRAN, *Johns Hopkins University*

THIS book was developed from an introductory course given to students in the various sciences. The only mathematical preparation assumed is a knowledge of elementary algebra.

The order in which the principal topics appear is indicated by the following series of headings: frequency distributions, the mean and standard distribution, sample-population relations, the normal distribution, confidence limits for the mean, the notion of a test of significance, the distribution of the estimated mean square s^2 and of F, the t-test of the mean and of the difference between two means, tolerance limits, control charts, analysis of variance for single and double classifications and the latin square, linear regression and correlation, analysis of covariance, chi-square tests for ratios and proportions, including binomial probability paper. This completes what is perhaps the basic part of the book. A number of miscellaneous topics are taken up in the remainder—some of the common power functions, the principal IBM machines, quick estimates from small samples, the sign test and tests based on runs, sequential analysis, sensitivity experiments.

The book is characterized by an emphasis on the inferential as distinct from the purely descriptive aspects of statistics, and by the early introduction of small-sample results. In these respects the authors follow the path taken previously by Snedecor. It might be said that they go further in this direction—notice the introduction of the F-test and of exact confidence limits for σ^2 before the introduction of the t-test. The inclusion of the quick but somewhat inefficient estimates of the mean and standard deviation based on order statistics is also of interest. The development of estimates of this type has progressed rapidly in recent years, and the day may not be distant when they are available for all standard elementary procedures. At this point the teachers of courses in elementary statistics will face the decision whether to substitute the new for the standard techniques. On this issue the authors take an orthodox approach by giving the standard methods, with the new ones mentioned, usually at a late stage in the book, as possible alternatives. This seems a reasonable and forward-looking stand in the present state of the subject.

Another feature is the wide range of topics covered, particularly in view of the small mathematical preparation demanded of the students. Inevitably, this means that numerous topics are discussed rather scantily, but the authors are successful in giving a lucid and succinct account of the important aspects of a topic. Individual teachers will doubtless wish to omit a number of topics in order to deal more fully with others. It is perhaps worth quoting the authors' treatment of degrees of freedom, from which they are determined to stand no nonsense. When the definition of s^2 is first intro-

duced (p. 20), they state, "The division by $N-1$ in place of N we shall subsequently find is necessary in order to use s^2 for the various functions it will perform in a number of statistical methods." This is not too informative, but having listened to long-winded and, it seemed to me, incorrect accounts of the reasons for using $N-1$, I have some sympathy with this high-handed attitude.

The book is noteworthy for its collection of tables, which extend over some 70 pages. In addition to the standard material, these include significance levels for some of the quick approximate tests, critical values of the sign test and of numbers of runs, the power function of the analysis of variance, and tolerance factors for the normal distribution.

The exposition is sound and clear, except for a few sentences which struck me as ambiguous. The authors do not attempt the derivation of any but the simplest results, but they do state the results forthrightly and correctly, with an explanation of the assumptions involved. There is a large collection of discussion questions and of problems to be worked by the students. In the great majority of the numerical illustrations, both in the text and the problems, the data appear to have been manufactured by the authors. This always appears regrettable to an applied statistician, though it is easier to complain about than to remedy, since the assembly of a wide range of actual examples which contain exactly the features desired is a most laborious task. Moreover, as the authors point out, their book is intended for a course which is a general introduction for students who will later study the applications of statistics to their particular fields.

The book is a competent piece of work, which I commend to the attention of beginning students and of teachers who give general introductory courses.

Introductory Statistics. *M. H. Quenouille* (Lecturer in Statistics, University of Aberdeen). London: Butterworth-Springer Ltd., 1950. Pp. xii, 248. 30s.

WILLIAM W. WRIGHT, *F. W. Dodge Corporation*

BRIDGMAN once remarked that statistical method is used either to conceal actual ignorance, or to smooth out details of physical complications.[1] In either event, statistics is now firmly wedded to a wide variety of disciplines from mathematical physics to market analysis. Interest in fundamentals is booming. Witness, for instance, the publication since 1946, of no fewer than twenty books bearing such titles as *Elements of*, *First Course in*, or *Introduction to Statistics*. Among these is *Introductory Statistics* by Maurice Henri Quenouille. Quenouille considers that, "an elementary non-mathematical manual on statistics could serve the two purposes of supplying students of statistics with a text-book, and of helping research workers to a fuller understanding of the methods used in the analysis of their results." His choice of illustrations and exercise material gives evidence of wide

[1] Bridgman, P. W., *The Logic of Modern Physics* (The Macmillan Company, 1927), page 205.

acquaintance with applied statistics. It also shows up some of his biases. For example, out of 102 problems scattered through the text, 53 (more than half) draw upon agricultural and biological data; 33 utilize material from the fields of anthropology, economics, psychology and sociology. Four relate to astronomy and meteorology. The remaining twelve are mathematical exercises. Quenouille, incidentally, justifies his selection of illustrative material on the interesting premise that "biological data are free from many of the complicating considerations, which often occur in economic and psychological data."

The range of topics considered by Quenouille is reminiscent of G. W. Snedecor's *Statistical Methods*. There, however, the comparison ends. The text consists of nine chapters each divided into two parts. The first part treats the subject matter on an elementary level; the second explores some of the theory and more advanced applications. Chapter 1 considers the organization of data, and measures of central tendency and dispersion. Chapter 2 gives a brief outline of frequency distributions and curves. Chapter 3 introduces the problem of comparing two samples. Quenouille uses as his main test procedures the variance ratio and the standard error of the difference between two means. Chapter 4 generalizes the type of problem considered in Chapter 3 to N samples (the analysis of variance) and explores applications of randomized blocks, Latin squares, confounding, and related matters. Chapter 5 is devoted to the χ^2-test. Chapter 6 reviews the elements of regression theory, and certain aspects of simple, partial, and multiple correlation. Procedures are also given for testing the significance of estimates of sample correlation coefficients. Chapter 7 discusses the reliability of regression coefficients, the analysis of covariance, and non-orthogonal comparisons. Chapter 8 deals with techniques for normalizing data under specific conditions, including the square root, $\sin^{-1}\sqrt{p}$, $\sinh^{-1}\sqrt{bx}$, and logarithmic transformations. Chapter 9 reviews the elements of random, stratified, systematic, and multi-stage sampling. Finally there is an appendix containing tables of the normal, F, and χ^2 distributions, logarithms, $\sin^{-1}\sqrt{x}$ and $\sinh^{-1}\sqrt{x}$, random numbers, and square roots.

Despite this wealth of material, the book is not entirely successful. It is too condensed for accuracy, elegance, or clarity. It reads in places like the transcript of a blackboard lecture—without the blackboard.

Often it fails to crystallize operationally important concepts. To illustrate, there are no useful definitions of population, sample, attribute, variable, random, or estimate. Other concepts are neither adroitly nor efficiently stated. This is particularly true of statements about the normal distribution, whose equation, incidentally, is never actually stated, although the less important Poisson distribution does appear. Somewhat the same criticism applies to the treatment of test procedures. The t-test is illustrated but not stated explicitly. Also, a student might well find the text inadequate on the subject of the number of degrees of freedom in analysis of variance. Even more important is that a novice could easily become confused between popu-

lation parameters and sample estimates of them. Occasionally Quenouille takes a position different from that of other statisticians. One instance is his insistence on computing the mean deviation from the mean rather than the median (p. 5), a practice which Cramér states "is not to be recommended."[2] Quenouille also tends to use "spread" in place of dispersion, generally, or standard deviation, specifically (p. 6). Finally, he collides with Yates (p. 205) on the question of estimating sample variances in systematic samples.[3]

Perhaps the most serious failing is in slurring over the nature of the statistical process. For statistics, as a discipline, is part of scientific methodology. Statistical data are not God-given. Instead three well-defined steps are necessary in fact-gathering operations. The first step consists in analyzing the problem, specifying the data required and setting-up procedures for collecting them. The second is the actual collection of data, using these procedures. The third involves making estimates from the data, and, where necessary, computing the sampling errors involved. Only after all three steps is the investigator in a position to draw inferences and test hypotheses. For a novice such precepts can be most helpful. The author never quite gets them across.

Quenouille does succeed in keeping to a minimum his mathematical demands on the reader. This, however, is not done without penalty. Too often, many words replace a few symbols, at the cost of both clarity and exactness. Karl Pearson once wrote of pure mathematics: "Since the concepts formed are in general narrowly defined and free from the infinite complexity . . . of perception, we are able to reason with great preciseness."[4] There may be a lesson there for all statisticians and, especially, for those writing text books.

Tables to Facilitate Sequential *t*-Tests. *National Bureau of Standards Applied Mathematics Series* 7. Washington: U. S. Government Printing Office, 1951. Pp. xix, 82. 45 cents.

THESE tables were prepared during the war for the Statistical Research Group, Columbia University in order to make sequential analysis practical for certain *t*-tests, that is, tests of hypotheses concerning a normal distribution when the standard deviation is unknown.

There are at least two kinds of sequential *t*-test, with two distinct types under each:

(1) Testing the mean of a normal population
  (a) against a one-sided alternative; i.e., where the tested value (null hypothesis) is to be rejected if the true mean appears to differ significantly in a specified direction.
  (b) against a two-sided alternative; i.e., where the null hypothesis is

[2] Cramér, Harald, *Mathematical Methods of Statistics* (Princeton University Press, 1946), page 181.
[3] Yates, F. "Systematic Sampling," *Philosophical Transaction of the Royal Society of London*, Series A, 241, 1948, page 345.
[4] Pearson, Karl, *The Grammar of Science* (J. M. Dent & Son Ltd., 1937), page 324.

to be rejected if the true mean appears to differ significantly, regardless of the direction of the apparent difference.

(2) Testing the proportion of a normal population exceeding a given value
   (a) against a one-sided alternative.
   (b) against a two-sided alternative.

Cases (1a) and (1b) are actually special instances of (2a) and (2b), respectively, with the tested proportion equal to one-half. These tables are intended only for case (1b), and then only when the alternative hypothesis is symmetrical with the null hypothesis, in the sense that amount of deviation which would require rejection of the null hypothesis is the same in both directions.

In considering a rather complicated procedure proposed for this case by the late Abraham Wald, it was noticed by Kenneth J. Arnold and the late Henry Goldberg that the result could be expressed in terms of a confluent hypergeometric function. (See also S. Rushton, "On a Sequential *t*-test," *Biometrika*, Vol. 37, 1950, pp. 326–333.) The tables here published were accordingly computed by the Mathematical Tables Project, then, like the Statistical Research Group, under the direction of the Applied Mathematics Panel of the Office of Scientific Research and Development.

An excellent introduction by Arnold describes and compares two slightly different sequential procedures for the problem considered, gives clear directions for using the tables, including an example, discusses interpolation to risks other than those covered and extrapolation to sample sizes beyond those covered, and concludes with a list of eight references.

There is an obvious typographical error in formula (15) of the Introduction: Omission of brackets around a sum makes what should be the square of a sum read as a sum of squares, thereby causing a ratio that should be a random variable to be identically 1.

The tables will be invaluable in making two-sided sequential tests of means of normal populations with unknown standard deviations.

W. A. W.

The Art of Asking Questions. *Stanley L. Payne.* Princeton: Princeton University Press, 1951. Pp. xiv, 249. $3.75.

M. Brewster Smith, *Vassar College*

This is a book on question-wording as an art. The text underlying Payne's sermon—a sermon delivered with engaging levity—may be found in a passage that he quotes from Stouffer and his collaborators in the *American Soldier:*

> To many who worked in the Research Branch it soon became evident that error or bias attributable to sampling and to methods of questionnaire administration were relatively small as compared with other types of variation—especially variation attributable to different ways of wording questions.

The task that Payne sets himself is to codify some of the "know-how"

accumulated in the polling profession in regard to the problem of framing single questions on matters of information or opinion. Within these self-imposed limits he has done a thoroughly creditable job that should be widely useful, especially to the neophyte but also to the seasoned practitioner.

Payne properly begins with the problem of defining and specifying the issue at which a question is to be directed. He then examines the major types of questions—free-answer, dichotomous, and multiple-choice—for their peculiar advantages and difficulties. The novice should be considerably helped by the wealth of example with which he illustrates how the various types of questions may be adapted to special purposes, and highlights the more common pitfalls. Considerations of rapport, clarity, and bias as they affect question wording occupy several following chapters. Payne selects from previously published word-lists an alphabetical list of 1,000 familiar words, which he keys to a detailed discussion of 89 "problem words." The phrasing of questions for oral reading is next considered. In a chapter from which the novice should learn much, the reader is finally taken through the trial-and-error process of wording a particular question. Starting with a preliminary formulation of the issue,

> Which do you prefer—dichotomous or open questions?

Payne leads the reader through its fuller specification,

> Which do you prefer to be asked in opinion surveys—dichotomous, multiple-choice, or open questions?

to a final wording:

> Which questions did you like best—those stating two answers to decide between, those stating more than two answers, or those leaving the answer for you to state?

It takes him 40 steps to get there. His admonitions are summarized at the end of the book in a convenient hundred-item check list.

Payne's references to the literature of opinion research are by way of vivid and instructive example rather than of documentation. This is a book of wisdom, not of the hard-earned fact that is slowly emerging from current programs of methodological research. But in all fairness, one must note that many of his cautions, once stated, are too obvious to warrant research, even though they are easy to neglect in practice. Laudably free from dogmatism, the book continually stresses the need for experimentation by split-ballot to resolve uncertainties about question-wording.

Payne makes amply clear his understanding that commitment to a single way of putting a question inevitably involves bias one way or another. On these grounds, the methodologically fastidious may eschew the single question in favor of rational scaling procedures. Yet even from such a position, to which the author gives no recognition, most of the problems with which he deals remain to be coped with in the wording of scale items. Statistical checks on the adequacy of a total scale do not supplant the sort of wisdom represented in this book.

Econometrics. *Dr. J. Tinbergen.* Translated from the Dutch by *Dr. H. Rijken van Olst.* Philadelphia: The Blakiston Company, 1951. Pp. xii, 258. $4.50.

M. G. KENDALL, *London School of Economics*

ECONOMETRICS is a science with a name, a society, several journals, an extensive literature, perhaps even a mystique, but with practically no text-books. C. F. Roos's "Dynamic Economics" of 1934 and H. T. Davis's "Theory of Econometrics" of 1941, together with a few books on mathematical economics, come to mind but they were written before the developments in structural analysis of the last decade, about which we are still waiting for an introductory account from the Cowles Commission. A contribution from Tinbergen is therefore greatly to be welcomed, particularly in an excellent English translation. Tinbergen is himself a distinguished writer in the econometric field and what he has to say concerning the subject has considerable authority.

This book has previously existed in a Dutch version which consisted of three parts: Introduction (the relationship of Econometrics to economics and mathematics), The Working Methods of Econometrics (mathematical formulation and statistical testing), and The Results of Econometric Research (psychic reactions, technical relations, the reactions of business life and the functioning of economic systems). The present English version has been revised and enlarged and there has been added a fourth part on Economic Policy together with an Appendix comprising three papers already published in econometric journals: on the use of correlation principles and long-term foreign trade elasticities. There is also a good index.

The main part of the book is more a survey, such as might be attempted in a lecture course to students of economics, than a text-book. In 190 pages with diagrams and a little elementary algebra, Tinbergen gives an account of what the econometrician tries to do and how he tries to do it. The statistician will probably find that the book raises for him more intellectual difficulties than it resolves. That, however, is hardly Tinbergen's fault but is inherent in the whole subject.

To the statistician who has been trained in the stricter scientific disciplines such as physics, economics is apt to appear an irritating mixture of unverifiable hypotheses and unverified conclusions. It begins with the explicit object of explaining economic phenomena from non-economic causes, which carries us at once into the mysteries of individual motivation and appears to try to explain observable from non-observable events. Its hypotheses are sometimes very vaguely stated. Literary economics at least tries to avoid a mathematical formulation of its arguments and hence discards one of the most powerful methods of clarification open to it. (As Tinbergen points out, the only linear equations which can be solved in literary terms are those with a diagonal matrix.) Until recently it tried to solve its problems in pieces instead of studying complete systems. And when it does formulate equations it does so sometimes in a manner which looks highly suspicious to a mathematician. For instance, on page 31 Tinbergen writes an equation and

continues "This supply equation must not be understood in the strict Walrasian sense of an equation, which determines the behaviour of the supply independently of the demand." That an equation can be true in a Walrasian sense but apparently untrue in another sense may well strike a mathematician as something that can only be true in a Pickwickian sense.

For these and other reasons econometrics is in a somewhat unsatisfactory state as compared with other sciences which make an extensive use of statistical techniques. It is generally pleaded that one should not shoot the economist because he is doing his best. One may, however, be permitted to fire an occasional shot in his direction to call him to a proper sense of his responsibilities. It is to be hoped that in future statisticians will do so and will turn more serious attention to the economic field than they have seemed inclined to do in the past. It may then be possible to produce a working manual for the statistical economist or for the economic statistician. In the meantime, Dr. Tinbergen's book is a useful introductory bird's-eye view of the field.

The Extent of Enterprise Monopoly in the United States, 1899–1939. *G. Warren Nutter.* Chicago: University of Chicago Press, 1951. Pp. xvi, 196. $4.00. Paper.

CLAIR WILCOX, *Swarthmore College*

IT is frequently asserted that competition has steadily declined during the past half-century, that monopoly has grown in extent, that it now characterizes most markets in the United States, and that the economies of large-scale production make its persistence inevitable. But these assertions are usually supported, if at all, by little more than a description of notorious cases and a listing of particular instances. They find no basis in actual measurements of the growth and extent of monopoly or its relative importance in the economy. It is to the task of measurement that Mr. Nutter turns his attention in the present study.

In comparing the extent of monopoly in the late 'thirties with that obtaining at the turn of the century, the author identifies as monopolistic all those industries in which the four largest firms produced half or more of the output, and measures the relative extent of monopoly by estimating the share of national income originating in these industries. On this basis he finds, in the most significant of his comparisons, that monopoly accounted for 19 percent of the income originating in non-governmental production in 1899 and for only 15 per cent of such income in 1937. The relative importance of monopoly, during this period, increased only in the case of financial enterprises; it remained the same in agriculture, services, trade, contract construction, and public utilities; it declined in manufacturing, mining, transportation, and communications.

In measuring the extent of monopoly in the late 'thirties, the author employs a more inclusive criterion, identifying as monopolistic all the industries that were so characterized by this reviewer in T.N.E.C. Monograph

No. 21, plus all other industries where Means found four firms producing half or more of the output in 1935, plus all other products where Crowder found four firms accounting for three-fourths or more of the output in 1937. Here, again, he measures the extent of monopoly by estimating the share of income originating in monopolized industries. On this basis, he finds that 74 per cent of non-governmental income originated in areas that were competitive and only 26 per cent in those that were monopolized.

In an effort to appraise the significance of economies of scale as a cause of monopoly in manufacturing, the author works out Lorenz distributions showing the cumulative percentages of output produced by cumulative percentages of establishments in the years 1904 and 1939. While his computations reveal some tendency toward increasing concentration, they afford slight support for the contention that relatively larger plants have been steadily displacing relatively smaller ones. The curves for the two years are virtually identical. The top tenth of the establishments accounted for 75.5 per cent of the output in 1904, 78.5 per cent in 1919, and 78.2 per cent in 1939.

Mr. Nutter has done his work with considerable care. He gives his reasons for defining monopoly largely in terms of market structure rather than market behavior and for rejecting as impractical the employment of theoretically superior criteria that run in terms of elasticities of demand. He presents his basic data in an extensive appendix, points out the limitations of of the data relating to market structures, and discusses the bias that may exist in the estimates of income produced. He recognizes that the extent of monopoly may here be understated or overstated as a result of arbitrary definition, errors in classification, inaccuracies in measurement, and exclusion from consideration of labor organizations, pure cartels, and competing imports.

Mr. Nutter closes with a warning that the findings presented in his study must be interpreted with extreme caution. It is the extremely cautious conclusion of this reviewer that these findings, while they may not serve completely to demolish the contentions of the decline-of-competition school of economists, have placed upon it an overwhelming burden of proof.

Inflation in the United States, 1940–1948. *Lester V. Chandler* (Professor of Economics, Princeton University). New York: Harper and Brothers, 1951. Pp. xi, 402. $4.50.

E. S. Shaw, *Stanford University*

Professor Chandler's book, written under the auspices of the Merrill Foundation for the Advancement of Financial Knowledge, is an admirably lucid description and analysis of inflation in the United States during World War II and its aftermath. The book is written primarily for lay readers, but economists should enjoy the author's facile applications of income and employment theory to the context of inflation. In their profes-

sional capacity, statisticians who read the book should acquire a new sense of urgency regarding work for them to do on economic time series and on index numbers.

Blame for the disproportionate rise of money spending relative to output that inflated price indexes (the index for consumer goods by 72 per cent) is put principally on federal fiscal policy during the war years, though well over half of the general price advance took place after price controls were relaxed in 1946. Inadequate taxation permitted an excessive rise of disposable incomes and, during the war, of current savings. Federal debt policy permitted these savings to accumulate in liquid form and, in addition, put the money-creating facilities of the banking system at the disposal of the public on very cheap credit terms indeed. Such direct controls as price ceilings, rationing, and allocations dammed up the enormous stock of buying power during and shortly after hostilities but were removed prematurely. When they were lifted, private spending stimulated in some degree by the amassed liquidity flooded commodity markets and, with an assist from pressures by labor unions and other special-interest groups for autonomous increases in prices on their wares, converted suppressed inflation into the real thing.

Professor Chandler indicates explicitly how economic policy should be improved upon if and when we have another war to fight. Both personal and corporate income taxes, including taxes on excess profits, should be significantly higher, though they could not be put on a pay-as-you-go basis without gravely undermining production and allocation incentives. Compulsory lending to the Treasury should supplant the voluntary savings programs that featured World War II. Some monetary reforms, such as 100-per cent reserve requirements on deposit increments, would perhaps assure the Treasury of adequate funds without opening the gates wide to private borrowing from the banking system. Once hostilities are over, interest rates should be raised sooner and higher than in the recent inflation. Direct controls clearly should be used again, the author feels, since they evidently succeeded so well in damping inflation while they held. And they should not be taken off so hastily again but should be patiently endured so that spending cannot rise virtually unchecked while capacity limitations restrain the flow of real goods and services. If direct controls can be held, the various pressure-groups should be more likely to forego the excessive demands for increases in money income that contributed in a secondary way to the inflationary experience of the nineteen-forties. The author suggests that more adequate controls would be put into effect if public authorities were not haunted, as they were in World War II, by the fear of chronic depression in private demand.

Professor Chandler limits his statistical demonstrations largely to the familiar series that government agencies make available on aggregative phenomena. As all of us do on occasion, he makes daring and dubious defla-

tions of value data to obtain his estimates of change in real variables. With the tastes and talents of a broad audience in mind, he has avoided econometricizing and has not entered into technical disputes on, say, the merits of cost-of-living indexes under a regime of direct controls.

Fortunately for reviewers, Professor Chandler has left himself open to some reproaches. His volume, for example, may lack the stature of some classic studies of inflation, because it has so little to say, beyond the widows-and-orphans argument, about the real costs of inflation. He is far too neglectful of the inefficient allocations of resources that can and do result from direct controls and excess-profits taxes. There are minor points to fret about, such as the author's use of the harried word "dynamic" to characterize the forces that bear most of the responsibility for inflation, and such as his apparent surprise, when none is called for, over the price-inelasticity of *total* consumption as distinct from *relative* consumer demands for individual goods. Despite such matters as these, this book is a very good description of our recent object-lesson on how not to run a war on the economic front.

Behavior of Wage Rates during Business Cycles. *Daniel Creamer*, with the assistance of *Martin Bernstein*. Occasional Paper 34. New York: National Bureau of Economic Research, Inc., 1950. Pp. 66. $1.00. Paper.

WILLARD D. ARANT, *Swift & Company*

BOTH business cycle theory and wage theory contain hypotheses involving the price of labor which have been difficult to test empirically because of the absence of suitable wage rate data.

Average hourly earnings, more readily available, differ from wage rates in a number of ways, especially in their inclusion of payments related to output as well as those related to time. To estimate the importance of these differences, the authors have constructed seasonally adjusted monthly indexes of wage rates in "all manufacturing," and in each of nine manufacturing industries in the United States, 1920–31; and in Class I railroads, 1920–39. They utilize similar indexes for seven British manufacturing industries, 1920–39. These have been compared with average hourly earnings series covering, as nearly as possible, the same industries. The conclusions are properly qualified because of the existence of certain deficiencies in the data, most of which, however, do not seem likely to affect the timing of cyclical turning points, the main subject of inquiry. For industry aggregates and at the major turning points of business activity, the authors conclude that the turning points in average hourly earnings are a reliable indicator of those in wage rates. They are a somewhat less reliable indicator for industry subgroups. The authors believe also that cyclical amplitudes of wage rates and average hourly earnings were closely similar during the periods studied, although this judgment rests largely on circumstantial evidence. If true, this conclusion suggests that some recent writers may have exaggerated the

extent to which business operations are "administered" to cause variations in earnings within a given wage structure.

Analyses of the wage data and comparisons of them with the NBER reference cycle dates and other series yield the principal conclusions that wage rates typically turned a substantial number of months later than business activity and employment, and that there was usually less cyclical adjustment in the price of labor than in the price of factory output.

The authors warn that in the years following World War II neither wage rates nor average hourly earnings serves its former purpose in many industries because some of the "fringe" payments to labor are excluded from the statistics. The industry selected for a rather detailed explanation of the cost of fringe benefits, however, is the anthracite coal industry, whose statistics are not included in the study, while the reader is left without knowledge as to the importance of "fringes" or "working rules" in the railroad industry, which is included.

Business Confidence and Business Activity. A Case Study of the Recession of 1937. *Douglas A. Hayes.* Michigan Business Studies Vol. X, No. 5. Ann Arbor: University of Michigan Press, 1951. Pp. 130. Paper.

KENNETH D. ROOSE, *Oberlin College*

THE basic objective of this book is to answer two questions: "Why was the economy unable to attain full employment of productive resources in 1936–1937, and why did we plunge into the depression of early 1938?" (Preface). Hayes' answer to both questions is that the most important causal factor was the failure of business confidence, although he admits that there is "room for sincere differences of opinion" (Preface) as to what accounted for this failure.

The method utilized in the study is to examine in detail the movements of various important economic series for which monthly or annual time series are available. With the exception of two charts, one of which (xi) compares rates of change in manufacturing inventories with rates of change in shipments, the other of which (xiv) shows bond yields on a semi-log graph, and one table (xii) which purports to explain the level of residential construction on the basis of a multiple regression equation, the statistical analyses are confined to a comparison of the arithmetic changes in selected time series.

The strongest sections of the study would appear to be those dealing with the inventory boom, in which an attempt is made to distinguish between voluntary and involuntary accumulation of inventories, with the capital supply in which the widening spread in yields on bonds is cited as evidence of the reduced availability of capital to lesser-known companies, and with the failure of construction to reach recovery levels in the thirties which was attributed to the rising costs of building.

In general, Hayes appears to have made judicious use of his data. Never-

theless an important weakness is that the pointed conclusions in the Preface as to the causal responsibility of business confidence for the recession are not well substantiated in the study itself. In addition, a major criticism of the whole study must be his failure to use certain data which on occasion would add strength to his conclusions, and at other points would call into question his conclusions. Thus, to give a few examples, the entire study is devoted to a consideration of the relationships of business confidence to business activity, but one of the most important sources of data on these relationships, the special United States Senate *Hearings*[1] in the spring of 1938, was not noted or used by Hayes. His failure to refer to these hearings and certain other materials makes his conclusions about the causal relationship of business confidence to business activity less convincing.

Other examples abound. Chapter I is largely concerned with dating the downturn in 1937 and the revival in 1938. But this analysis is unnecessary since the National Bureau of Economic Research (unless Hayes wishes to criticize its conclusions) employed the 1937–1938 Recession and Revival as its model for demonstrating the problems encountered in dating peaks and troughs in business cycles.[2] Chapter II discusses profit and inventory trends and Hayes comments on the paucity of such data (p. 30 and pp. 45–46). Yet the comprehensive profit and inventory series of the National Bureau of Economic Research are not used.[3] As a final example, of the ten major studies on the period known to this reviewer, only *seven* are cited in Hayes' bibliography and four of these have inadequate reference if the complete position of the respective authors on the causal developments is desired.

Partly because these and other materials which could be mentioned are lacking, Hayes is led into outright error at times and occasionally into a generalization which finds inadequate support in his own presentation but in fact is covered effectively in the literature. Numerous examples of this might be given but a few will suffice. There is considerable doubt that the New Deal consciously had accepted compensatory spending as its guide to fiscal policy as early as 1937 (p. 19 and p. 114).[4] Hayes concludes on the basis of only minor evidence that seasoned industrial companies could obtain funds by sale of common stock but other companies faced more difficulty (p. 72). This argument has been developed at length elsewhere, with supporting data.[5] The decline in U. S. Government bond prices in February 1937 is said to have initiated the break in corporate bond prices and finally

---

[1] U. S. Congress, Senate, Special Committee to Investigate Unemployment and Relief, *Hearings*, Senate Resolution 36, 75th Congress, Third Session, I (January 4–22, 1938), and II (February 28–April 8, 1938).

[2] Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (New York, National Bureau of Economic Research, 1946).

[3] Harold L. Barger, *Outlay and Income in the United States, 1921–1938* (New York, National Bureau of Economic Research, 1942).

[4] Sherwood Fine says this position was not reached until after the recession. *Public Spending and Postwar Economic Policy* (New York, Columbia University Press, 1944), pp. 117–119.

[5] Kenneth D. Roose, "Federal Reserve Policy and the Recession of 1937–1938," *The Review of Economics and Statistics*, XXXII (1950), pp. 179–182.

in stock prices (p. 86 and p. 124). But the reviewer has shown elsewhere that the first price decline was in short-term Treasury bills and notes in December 1936, followed by municipal bonds early in January 1937, corporate bonds in the middle of January 1937 and last of all U. S. government bonds in late February 1937.[6] Differences of opinion also exist as to whether utility companies actually tabled capital expansion plans because of hostile public policies (p. 116). Finally, Hayes has not presented data to justify the conclusion that profit rates and margins in 1936–1937 did not return to the levels of the mid-twenties (p. 120).

This reviewer regrets not so much any errors of fact or conclusion as that the analysis begun carefully in this study was not extended to the many unused sources of data.

Occupational Choice, an Approach to a General Theory. *Eli Ginzberg, Sol W. Ginsberg, Sidney Axelrad,* and *John L. Herma.* New York: Columbia University Press, 1951. Pp. viii, 271. $3.75.

HENRY B. MCDANIEL, *Stanford University*

MANY years ago, Professor Harry Kitson wrote a book on the subject of occupational choice. It was addressed to youth and bore the title, "How to Find the Right Vocation." Out of his experience in vocational counseling, Professor Kitson advised youth that choosing a vocation is a process involving a careful, rational assessment of the individual and his environment. He told youth how to go about this business of choosing an occupation. Dr. Ginzberg and his associates deal with the same subject, choosing an occupation, in an almost exactly opposite manner. They seek to find out from youth how occupational choices are made.

This is a research report. The manuscript is divided into five parts: Design of the Study, The Basic Process of Decision Making, Significant Groups, A General Theory of Occupational Choice, Policy.

The authors consider their work as exploratory. Their goal is the development of a general theory of occupational choice. This is obviously a long range goal. The present study is a beginning. It attempts to develop a method for gathering and ordering the data. The nature of the data to be sought is indicated by this statement: "But we believe that a meaningful approach must concentrate on the decisions which the individual makes with respect to his ultimate work during successive periods of his maturation. These decisions hold the key to an understanding of the process" (p. 25).

The design of the study follows this assumption. The data analyzed come from interviews with groups of youth at different age levels. Eight white males were interviewed at each of eight stages in the educational process, starting with the sixth grade, through the eighth, tenth, and twelfth grades; the freshman and senior years in college; first year graduate students and advanced graduate students. All were from Horace Mann-Lincoln

[6] *Ibid.*, p. 179.

School and Columbia University. All were from a relatively high socio-economic level and *were deemed to have freedom of choice.* This is a highly restricted sampling, the primary difference between groups being a difference in age and educational level. Certain additional data were obtained from a group of 17 boys from a settlement house and a group of 10 young women from Barnard College.

A moderately structured interview technique was used which sought to obtain formation on the individual's concept of self, of reality, and of key persons in his environment. Most interviews were recorded by a stenographer in the room but out of the line of vision of the subject. Interviewers found that most of their subjects talked freely about their educational and occupational problems. Topics under the "self" category included, in addition to the usual interests, capacities, and values questions, a careful consideration of "time perspective" queries. "Reality" questions concerned family, education, world of work, and life plan. Following the interview, the complete typescript was condensed to a single page summary. This summary then formed the basis of a group discussion among the four research collaborators. In this discussion analysis, these four, a sociologist, a psychiatrist, an economist, and a psychologist, sought to evaluate the subject's own statements about himself, the reality factors, and the key persons who were influencing his occupational choice process.

In this setting and from these data, the investigators examined the nature of occupational choice. Their findings, supported by digests and direct excerpts from the interview data, are reported in a number of chapters in this book. No attempt has been made to apply statistical methods of analysis to the data. The data have been synthesized and ordered qualitatively through group discussion. Among the several characteristics of occupational choice presented, two aspects will be discussed here. These are the division into time sequence periods and the general conclusions.

Occupational choice appears to be divided into three general time periods: fantasy choices, tentative choices, and realistic choices. Each of these has subdivisions. The period of fantasy choice spans roughly the period from age six to age eleven. Occupational choices during this period are based on activities that give pleasure. Barriers of either a personal or environmental nature are not considered. Choices change freely and frequently. The period of tentative choices ranges roughly from eleven through seventeen years. Educationally, this period spans the secondary school. The authors find that the period embraces four stages: the interest stage, the capacity stage, the value stage, and the transitional stage. The individual learns gradually that interests must be supported by capacities and that such reality considerations as employment opportunity, reward, and social utility must fit into his final synthesis. At the end of this period, the youth must begin to implement his choice, either through schooling beyond the compulsory period or through entry into work activities. The period of realistic choices has as its essential characteristic the consideration of external reality factors. The

tentative choice is tested by further exploration. For the group in this study, this process included the first year of college with new freedoms, new subjects of study, and new experiences with people. The outcome of this exploration is clarification of and commitment to the choice made.

The final conclusions reached in this study came as no surprise to one who reads the report, with its emphasis on sequence, order, and process. *Occupational choice is a development process.* It is not an event nor a single decision. *The process is largely irreversible.* Each succeeding step is based on the preceding one. Conditioning experiences cannot be relived. Refinement and specification can take place, but shifting to a different field is extravagantly wasteful of both time and experience. *The process ends in a compromise.* The final synthesis of determinants, interests, capacities, opportunities, values is a balance among all, a compromise. These three statements are the elements of the general theory of occupational choice which these investigators sought to discover.

To one experienced in the field of vocational counseling, these findings are not startling. They fit into and bring together the findings and observations of other workers. One could pick out statements throughout the report and question whether or not a finding from a small and restricted sample could be so generalized. One could point out the danger of over-simplification in the neat categories of three and four stages and aspects of this complicated and interrelated process. One could also question the reality of the separation of occupational decision making from other life problems. Yet, for purposes of study, analysis and even some atomization is necessary. If one reads all of this book, he finds that the writers have pointed out the limitations of their data, the restrictions with which the findings must be interpreted. Yet some sections appear to over-generalize. One does not find enough repetition of the condition "in terms of the data from this sample." Educators will need to be wary of trying to build a guidance program around the four developmental stages of the tentative choice period.

Much more could be said of the methodology of this research. This reviewer is impressed with the value of the team approach which brings several disciplines to bear on the whole process of planning and implementing. The findings always need further testing. The authors themselves suggest many topics for further research. The ideas discussed are of serious import to many people; parents and teachers as well as professional counselors.

Use of Factorial Design in Industrial Relations Research. *Clarence W. Nelson.* Research and Technical Report 6, Industrial Relations Center, University of Minnesota. Dubuque: William C. Brown Company, November, 1950. Pp. iii, 52. Paper.

W. S. Connor, Jr., *National Bureau of Standards*

The purpose of this Report is to acquaint persons who do research in industrial relations with the factorial design. The Report is essentially a

manual, giving detailed instructions on how to set up and analyze the design, with almost no discussion of statistical theory. As a manual the Report is well written and is to be recommended to industrial relations personnel.

It is pointed out in Part One that an m-factor design will give the same information on the factors as is obtained at greater cost from m non-factorial designs, and in addition, the factorial design furnishes information about the relationships between factors which is not available from the non-factorial designs. Part Two gives a detailed description of how to set up the design, calculate the various quantities needed for the analysis, and draw conclusions. The two-factor design which is used as an illustrative example in this part treats a problem familiar to industrial relations personnel. Twenty-four new employees are to be trained for a job, and the two factors to be studied are (i) method of training and (ii) instructor. It is desired to determine the best method, best instructor, and whether an instructor teaches better when using one method than when using another.

Part Three is a mathematical appendix, which might well have been omitted, and instead reference made to mathematical statistical treatments of the subject.

Several omissions in the Report are worth pointing out. As one step in the analysis it is recommended that "when components have been determined to be non-significant, their sums of squares and their degrees of freedom may be combined with those of the error component". . . . It should be cautioned that this step is not to be taken unless the experimenter has strong *a priori* grounds for believing a component to be non-significant.

The discussion of the F test does not state that the F test is strictly applicable only when the observations are normally distributed. For a discussion of the error made in using the F test when the data are not normal, one may consult reference 19 of the Classified Bibliography appended to the Report.

The Report makes no mention of the t test, which could be very useful in comparing any two methods or foremen.

But apart from these observations, the Report will have served admirably if it attracts the interest of persons in industrial relations research to the use of the factorial design, and indeed of other experimental designs in their work.

Studies of Sampling Techniques and Chemical Analysis of Vegetables. *Southern Cooperative Series Bulletin No. 10.* March 1951. Pp. 143.

BOYD HARSHBARGER, *Virginia Polytechnic Institute*

THE purpose of the bulletin as expressed by the authors "is to present the results of studies of sampling techniques and chemical analyses performed by the Southern Cooperative Group in connection with their studies of the nutritive value of certain vegetable crops." This Southern Cooperative Group is drawn from a number of southern Land Grant Colleges, and they

work together as a team on the determinators of nutritive values of vegetables. The organization has been successful in developing techniques and coordination, and in supplementing findings of the research of its members. How much it has retarded individual initiative is open to question. It has been of great help to the technician sort of research. Some of the findings of this group have been used in the contents of the bulletin.

The bulletin will serve as an example of good statistical techniques and will encourage a better appreciation of their usefulness to the people in the field of nutrition. There is no attempt made to use new statistical methodology that is not commonly found in any good elementary statistical text.

The main contribution of this publication is the estimates given of the magnitude of certain errors and the use of these for estimating sample sizes needed in future experiments. Estimates of variance of the interaction for places by treatments or varieties are not used to their fullest advantage.

The bulletin gives pertinent and useful information including the Analysis of Variance Tables for the following:

1. Selection of Samples from the Field.
2. Holding of the Sample.
3. Preparation of Samples.
4. Size of Samples for Analysis.
5. Determination of Ascorbic Acid.
6. Determination of Carotene.
7. Determination of Riboflavin.
8. Determination of Calcium, Magnesium and Phosphorus.

It can be seen from these headings that the study is very comprehensive and brings together many ideas on sampling techniques and chemical analysis of vegetables that were formerly scattered through the literature. The bulletin should be read by anyone embarking on the chemical analysis of vegetables.

Monte Carlo Method. *National Bureau of Standards Applied Mathematics Series 12.* Washington: U. S. Government Printing Office, 1951. Pp. cii, 42. 30 cents.

THIS pamphlet consists of thirteen papers and a summary of discussion at a three-day conference on the Monte Carlo method sponsored by the Rand Corporation and the Bureau of Standards during the summer of 1951. Statisticians will be especially interested in the last four papers, by George W. Brown, Preston C. Hammer, George E. Forsythe, and John von Neumann, which deal with various methods of producing and testing random digits.

The Monte Carlo method has been discussed recently in this JOURNAL by Nicolas Metropolis and S. Ulam; see Vol. 44, 1949, pp. 335–41.

W. A. W.

PUBLICATIONS RECEIVED

Beglinger, J.-F. *Les Conditions Economiques de la Paix.* Collection de la Nouvelle Ecole de Lausanne. Librairie F. Rouge, Lausanne. Paper.

Birmingham, City of. *Annual Abstract of Statistics, Number One, 1931–1949.* City of Birmingham Statistical Office, England, 1951. 15 s.

The Brookings Institution, International Studies Group of. *Major Problems of U. S. Foreign Policy,* 1951–52. Washington, D. C., 1951. Paper, $1.50. Cloth, $3.00.

Broom, Halsey N. *Application of Statistical Methods to Factory Management Problems.* Bureau of Business Research, University of Texas, Austin, 1951. Paper. $1.50.

Cook, Robert C. *Human Fertility, the Modern Dilemma.* William Sloan, Associates, New York, 1951. $4.50.

Daval, R. and Guilbaud, G. T. *La Méthode Statistique; Première Partie: Statistique à un seul caractère.* Presses Universitaires de France, Paris, 1951. Paper.

Denmark, Copenhagen. *Det Statistiche Department, 1920–50.* Paper. Kr. 3.00.

Dressel, Paul L. and Schmid, John. *An Evaluation of the Test of General Educational Development.* American Council on Education, Washington, D. C., 1951. Paper. $1.00.

Fisher, Ernest M. *Urban Real Estate Markets, Characteristics and Financing.* National Bureau of Economic Research, New York, 1951. $3.00.

Hamburg, D. *Business Cycles.* The Macmillan Company, New York, 1951. $5.00.

Illinois, University of and Chicago Chapter of the A.S.A. *Business Application of Statistical Sampling Methods.* Proceedings of Conference May 25 and 26, 1950. Paper. $2.00.

India, Government of. *First Report of the National Income Committee,* April 1951. Minister of Finance, Calcutta. Paper. 2 s.

Jaffe, A. J. and Stewart, Charles D. *Manpower Resources and Utilization.* John Wiley and Sons, Inc., New York, 1951. $6.50.

Juran, J. M. *Quality Control Handbook.* McGraw-Hill Book Company, New York, 1951.

Levinson, Harold M. *Unionism, Wage Trends, and Income Distribution, 1914–1947.* University of Michigan Press, Ann Arbor, 1951. Paper.

Lorie, James H. and Roberts, Harry V. *Basic Methods of Marketing Research.* McGraw-Hill Book Company, New York, 1951. $6.00.

Love, Esther M. *Operating Results of Limited Price Variety Chains in 1950.* Harvard University Graduate School of Business Administration, Cambridge, July 1951. Paper. $2.00.

Mackay, Donald. *Hospital Morbidity Statistics, A Preliminary Study of In-patient Discharges.* H. M. Stationery Office, London, 1951. Paper. 3 s. 6 d.

Mills, James I. *A Proposed System for Classifying Manufacturing Concerns by Size.* U. S. Department of Commerce, National Production Authority, Washington, D. C., 1951. Paper.

Monroe Calculating Machine Company. *Quality Control, Some Monroe Short-Cut Statistical Methods.* Orange, N. J., 1951. Paper.

Nagell, Trygve. *Introduction to Number Theory.* John Wiley and Sons, Inc., New York, 1951. $5.00.

National Bureau of Standards. *Problems for the Numerical Analysis of the Future.* Government Printing Office, Washington, D. C., 1951. Paper. 20 cents.

National Opinion Research Center. *The Role of Unemployment Compensation in Maintaining Family Income and Expenditure in an Area of Critical Unemployment.* Survey No. 274, Report No. 43. University of Chicago, August 1951. Paper.

Nyblén, Göran. *The Problem of Summation in Economic Series.* Lund, 1951.

Quebec, Ministry of Trade and Commerce. *Statistical Yearbook, Quebec, 1950.*

Royal Statistical Society, Council of the. *The Teaching of Statistics in Schools.* Printed in Advance of the Journal of the Royal Statistical Society Series A (General) Vol. CXV, Part I, 1952. 1951. Paper. 1 s.

Schreier, O. and Sperner, E. *Introduction to Modern Algebra and Matric Theory.* Chelsea Publishing Company, New York, 1951.

Seattle, City of, Planning Commission. *The Population of Metropolitan Seattle, An Interim Report.* December 1950, Reprinted June 1951. Paper.

Sittig, J. and Feudenthall, H. *De Juiste Maat.* L. Staflen, Leiden, 1951. f 20.

Slavick, Fred. *The Operation of Sickness Benefit Plans in Collective Bargaining.* Princeton University, 1951. Paper. $2.50.

Smith, Kenneth. *The Malthusian Controversy.* Routledge and Kegan Paul, London, 1951. 30 s.

Snedecor, George W. *Everyday Statistics.* Wm. C. Brown Company, Dubuque, Iowa, 1950. Paper, spiral bound. $3.00.

Tableman, Betty. *Governmental Organization in Metropolitan Areas.* University of Michigan Press, Ann Arbor, 1951. Paper.

Thompson, Warren S. *Migration within Ohio, 1935–40; A Study in the Re-Distribution of Population.* Scripps Foundation for Research, Miami University, Oxford, Ohio, 1951. Paper.

Tibor, Scitovsky, Shaw, Edward and Tarshis, Lorie. *Mobilizing Resources for War.* McGraw-Hill Book Company, New York, 1951. $4.50.

Todd, John. *Table of Arctangents of Rational Numbers.* National Bureau of Standards, Washington D. C., 1951. $1.50.

Traweek, Stella. *The Use of the Punched-Card Method by an Apparel Specialty Store.*

———. *The Use of the Punched-Card Method by a Clothing Manufacturer.*

———. *The Use of the Punched-Card Method by a Motor Freight Line.*

———. *The Use of the Punched-Card Method by a Retail Grocery Chain.* Bureau of Business Research, University of Texas, Austin, 1951. Paper. 25 cents each.

U. S. Bureau of Labor Statistics. *Publications of the Bureau of Labor Statistics, January–June 1951.* Government Printing Office, Washington, D. C., 1951. Paper.

U. S. Department of Commerce, Bureau of Foreign and Domestic Commerce. *National Income Edition, 1951 Supplement to the Survey of Current Business.* Government Printing Office, Washington, D. C., 1951. Paper. $1.00.

U. S. House of Representatives. *Consumer's Price Index—Hearings before the Subcommittee of the Committee on Education and Labor.* Government Printing Office, Washington, D. C., 1951. Paper.

World Health Organization. *Annual Epidemiological and Vital Statistics, 1939–1946. Part I, Vital Statistics and Causes of Death.* Geneva. Paper. $4.00.